2005

Dynamical systems with time delay

Norma Ortiz
Louisiana State University and Agricultural and Mechanical College, ortiz@math.lsu.edu

DYNAMICAL SYSTEMS WITH TIME DELAY

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Mathematics

by
Norma Ortiz
B.S. in Computer Sci., University of Louisiana, Lafayette, 1996
M.S., Louisiana State University, 2001
May 2005

Acknowledgments

This dissertation would not be possible without several contributions. First, I would like to thank my advisor, Professor Peter R. Wolenski, for his guidance, patience, and enthusiasm about mathematics. From the beginning, when I attended his course in Differential Equations, he was very supportive and available to discuss mathematics with me. I would also like to thank the members of my Advising Committee for their support and direction. In particular, Professor Neubrander was a great source of direction and encouragement as I made decisions along the way. In addition, special thanks are extended to Professor Oxley who provided guidance and support in the area of teaching and to Professor Shipman who was always available to discuss mathematics and share his experiences in the mathematical field. It is a pleasure to thank the staff in the mathematics department for providing me with a great working environment. I especially thank Mr. Anthony Picado and Mr. Jeff Sheldon.

I would especially like to thank my family and friends for their support and encouragement throughout the past few years. I would not have been able to finish this work without you. Finally, I thank God for giving me life, direction, wisdom, and for doing in my life greater things than I ever imagined possible.

Table of Contents

Abstract

In this dissertation, we study necessary conditions and weak invariance properties of dynamical systems with time delay. A number of results have been obtained recently that refine necessary conditions of optimal solutions for nonsmooth dynamical systems without time delay. In this dissertation, we examine the extension of some of these results to problems with time delay. In particular, we study the generalized problem of Bolza with the addition of delay in the state and velocity variables and refer to this problem as the Neutral Problem of Bolza. We consider the relationship between the generalized problem of Bolza with time delay and control systems, establish existence of solutions for the Neutral Problem of Bolza, and use a "decoupling" technique introduced by Clarke [2] to derive necessary conditions of Hamiltonian and Euler-Lagrange type for this problem. We also apply the same methods to the generalized problem of Bolza with time delay in the state variable only and compare the results obtained in this case with the results obtained in the neutral case. Furthermore, we study the system (S, F) involving a closed set S and a delayed autonomous multifunction $F(x(t), x(t-\Delta))$. Under suitable hypotheses, we provide a characterization of weak invariant properties for F in terms of the lower hamiltonian.

Introduction

In this dissertation we study dynamical systems with time delay. Systems with time delay are of interest since it has been shown that when modeling systems in engineering, finance, and others, there is almost always a time delay [9, 10]. More specifically, we study necessary conditions for optimal solutions and weak invariance properties of nonsmooth dynamical systems with time delay. In our study of necessary conditions, we study the Neutral problem of Bolza (NPB), which consists of the generalized problem of Bolza introduced by Rockafellar [24] with added dependence on time delay in the state and velocity variables. We consider the relationship between the generalized problem of Bolza with time delay and general control systems, establish existence of solutions for NPB, and use a "decoupling" technique introduced by Clarke [2] to derive necessary conditions of Hamiltonian and Euler-Lagrange type.

Earlier work in the calculus of variations involving neutral systems, [9],[12],[30], established existence and necessary conditions for the problem min $J(\cdot)$, where $J(\cdot)$ is an integral functional of the form

$$J(y) := \int_a^b f[t, y(t-\tau), y(t), \dot{y}(t-\tau), \dot{y}(t)]\, dt.$$

In this problem, minimization was considered over the class of piecewise smooth functions $y(\cdot)$ and the function f was assumed to be smooth. A relationship between the calculus of variations and control systems was established by Rockafellar who introduced the generalized problem of Bolza [28] as a means to study general control systems. The generalized problem of Bolza uses the technique of infinite penalization [4], [13], [14], [31] in order to incorporate constraints. Given the use of infinite penalization, smoothness is no longer a viable assumption for the Lagrangian involved.

A number of results have been obtained recently that refine necessary conditions for non-smooth dynamic optimization systems without time delay. Particularly noteworthy are papers by Rockafellar [24] and Clarke [2] that establish a non-smooth unified version of the Euler-Lagrange and Hamiltonian necessary conditions for optimal solutions of dynamic systems. Further results by Loewen and Rockafellar, [13, 14], demonstrate the existence of this unification in a more refined form. Both of these results were obtained in the non-smooth case for systems without time delay. Furthermore, in his recent work [4], Clarke shows a unification of the Euler-Lagrange and Weistrass conditions, again, for non-smooth systems without time delay. These results are of much importance and in our research we consider if they continue to hold for systems with time delay in the state and velocity variables.

Previous research to obtain necessary conditions for the case of state time delay problems includes the work of Mordukhovich and Trubnik, [16]. In [16], a method of discrete approximations is used to obtain the Euler-Lagrange necessary conditions for the Bolza problem and the unified Euler-Lagrange and Hamiltonian necessary conditions for the Mayer problem, a problem similar to the Bolza problem. Some optimal control problems of neutral type have been studied in recent work by Mordukhovich and Wang [17],[18]. Their research is based on a technique that involves discrete approximations. In [17], necessary conditions are derived for the Mayer problem involving differential inclusions of the neutral type and [18] treats the Bolza problem involving a neutral differential inclusion as a constraint. Our problem formulation and results are different since we incorporate the delayed derivative in the Lagrangian itself, instead of treating it as a constraint, and we optimize over a different class of functions.

Our study of weak invariance is brief and presents an opportunity for further study. We provide a characterization for weak invariance properties of a system (S, F) consisting of a closed set S and a delayed autonomous multifunction $F(x(t), x(t-\Delta))$ for a fixed real number $\Delta > 0$. Weak invariance properties are an important aspect of dynamical systems and are related to the Hamilton Jacobi equation, which can be used to give a sufficient condition for optimality of dynamical systems. Weak invariance properties for delayed multifunctions have been previously characterized in terms of the Bouligand contingent cone by Haddad [11]. We characterize weak invariant properties in terms of the lower hamiltonian and arrive at our results independently.

Chapter 1
Preliminaries

The basic terminology used in this dissertation will follow the textbook *Nonsmooth Analysis and Control Theory* [3]. A summary of the main concepts used in this dissertation follows. Throughout, X is a real Hilbert space and S is a non-empty closed subset of X. The functions considered will be extended real valued with range $(-\infty, \infty]$. For any given vectors $\alpha, \beta, \gamma \in X$ and any real number c, we denote $\langle \alpha, \beta \rangle$ as the *inner product* of α and β, which satisfies the following properties:

1. $\langle \alpha + \beta, \gamma \rangle = \langle \alpha, \gamma \rangle + \langle \beta, \gamma \rangle$;
2. $\langle c\alpha, \beta \rangle = c\langle \alpha, \beta \rangle$;
3. $\langle \beta, \alpha \rangle = \langle \alpha, \beta \rangle$;
4. $\langle \alpha, \alpha \rangle > 0$ if $\alpha \neq 0$.

The norm in the Hilbert space X is given by $\|\alpha\| = |\langle \alpha, \alpha \rangle|$ and satisfies the following.

1. $\|c\alpha\| = |c|\|\alpha\|$;
2. $\|\alpha\| > 0$ for $\alpha \neq 0$;
3. $|\langle \alpha, \beta \rangle| \leq \|\alpha\|\|\beta\|$;
4. $\|\alpha + \beta\| \leq \|\alpha\| + \|\beta\|$.

1.1 Basic Concepts in Nonsmooth Analysis

The definitions in this section are adapted from Chapter 1 of *Nonsmooth Analysis and Control Theory* [3]. A function $f : X \mapsto (-\infty, \infty]$ is said to be *lower semi-continuous* at a point x provided that $\liminf_{x' \to x} f(x') \geq f(x)$. We say that f

is an element of $\mathcal{F}$ if f is lower semi-continuous in X and its effective domain $dom f := \{x \in X : f(x) < \infty\}$ is not empty; the latter means that there is a point x for which $f(x)$ has finite value.

Suppose x is a point not lying in S. If there is a point $s \in S$ whose distance to x is minimal then we say that s is a *projection* of x onto S. We denote the set of all closest points by $proj_S(x)$. The distance function $d_S(\cdot)$ measures the shortest distance from a point x to the set S; it is defined by $d_S(x) = \inf\{\|x - s\| : s \in S\}$.

For $s \in S$, the *proximal normal cone* to S at s is denoted by $N_S^P(s)$. A proximal normal vector at the point s defines the direction of perpendicular departure from the set S. Formally, a vector $\zeta \in N_S^P(s)$ if and only if there exists a number $\sigma \geq 0$ such that

$$\langle \zeta, s' - s \rangle \leq \sigma \|s' - s\|^2, \ \forall s' \in S.$$

The *limiting normal cone* to S at $x \in S$ is denoted and defined by $N_S^L(x) := \left\{w - \lim \zeta_i : \zeta_i \in N_S^P(x_i), x_i \xrightarrow{S} x\right\}$. Here, $x_i \xrightarrow{S} x$ means that $x_i \to x$ where $x_i \in S$ for all i and $w - \lim$ denotes the weak limit.

A vector $\zeta \in X$ is called a *proximal subgradient* of a lower semi-continuous function f at $x \in dom\, f$, denoted as $\partial_P f(x)$, provided that

$$(\zeta, -1) \in N_{epif}^P(x, f(x)).$$

Given $f \in \mathcal{F}$ and $x \in dom\, f$ then $\zeta \in \partial_P f(x)$ if and only if there exist positive numbers σ and η such that

$$f(y) \geq f(x) + \langle \zeta, y - x \rangle - \sigma \|y - x\|^2 \ \forall y \in B(y; \eta). \tag{1.1}$$

We refer to the above inequality (1.1) as the proximal subgradient inequality ([3],Thrm. 2.5, pp. 33-34). In addition to the proximal subgradient, we will refer to the *limiting subgradient* of a lower semi-continuous function. The limiting subgradient (or subdifferential) is the set of all vectors ζ that can be expressed as the weak limit of a sequence $\{\zeta_i\}$, where $\zeta_i \in \partial_P f(x_i)$ for each i, and where $x_i \to x, f(x_i) \to f(x)$. It is formally denoted and defined as

$$\partial_L f(x) := \left\{ w - \lim \zeta_i : \zeta_i \in \partial_P f(x_i),\, x_i \xrightarrow{f} x \right\}.$$

An important result in the study of non-smooth analysis is known as the *Density Theorem* ([3], Thrm. 3.1, pp. 39-42). This theorem states that assuming $f \in \mathcal{F}$, for any $x_0 \in dom f$, when given an $\epsilon > 0$, we are guaranteed (1) the existence of a point $y \in x_0 + \epsilon B$ with $\partial_P f(y) \neq \emptyset$ and (2) the $dom(\partial_P f)$ is dense in $dom f$.

In his book *Optimization and Noonsmooth Analysis*, Clarke introduces the *generalized directional derivative* that characterizes a *generalized subgradient* for locally Lipschitz functions mapping a Banach space into the set of real numbers ([1], pp. 25-28). For a locally Lipschitz function $g : Y \to I\!R$ where Y is a Banach space, the Clarke generalized directional derivative g° and generalized gradient ∂_C are given by

$$g^\circ(\bar{x};\mu) = \limsup_{x \to \bar{x},\, \lambda \downarrow 0} \frac{g(x + \lambda\mu) - g(x)}{\lambda}$$
$$\partial_C g(\bar{x}) = \{p : \langle p, \mu\rangle \le g^\circ(\bar{x};\mu)\ \forall \mu \in I\!R^n\}.$$

1.2 Piecewise Smooth and Absolutely Continuous Functions

A function $x(\cdot)$ that maps a closed interval $[a, b]$ into $I\!R^n$ is said to be *piecewise smooth* (pws) if the following hold: x is continuous in $[a, b]$, there exist points t_i so that $a = t_0 < t_1 < \cdots < t_N = b$, $\dot{x}(t)$ exists at all $t \in [a, b]\backslash\{t_i\}_{i=1}^N$, $\dot{x}(\cdot)$ is continuous at each open subinterval (t_i, t_{i+1}), and $\dot{x}(\cdot)$ has one-sided limits at all $t \in [a, b]$. Notice that $x(\cdot)$ is a piecewise smooth function if and only if there exists a function $v(\cdot)$ so that

$$x(t) = x(a) + \int_a^t v(s)\,ds \tag{1.2}$$

and $v(\cdot)$ is piecewise continuous. Also, notice that *absolutely continuous* functions, or *arcs* are functions for which (1.2) holds with $v(\cdot)$ integrable but not necessarily piecewise continuous.

1.3 Optimal Solutions and Necessary Conditions

Consider the calculus of variations problem, $\min J(\cdot)$, where $J(\cdot)$ is an integral functional of the form

$$J(y) := \int_a^b L[t, y(t-\tau), y(t), \dot{y}(t-\tau), \dot{y}(t)]\,dt$$

with the given constraints $y(a) = A$, $y(b) = B$. Here A, B are each in $\mathbb{R}^n$ and we minimize over all absolutely continuous functions y. When an optimal solution is found it may be categorized as either a "weak"or a "strong" solution. A solution is termed "weak" when there exists a $\delta > 0$ such that $J(x) \leq J(y)$ for all arcs $y(\cdot)$ satisfying

$$\sup_{a\leq t\leq b} \left| x(t) - y(t) \right| \leq \delta \tag{1.3}$$

and

$$\sup_{a\leq t\leq b} \left| \dot{x}(t) - \dot{y}(t) \right| \leq \delta, \tag{1.4}$$

where $x(\cdot) : [a, b] \to \mathbb{R}^n$. It is understood that if either $\dot{y}(t)$ or $\dot{x}(t)$ does not exist for some some t in $[a, b]$, then we replace it with its left or right limit. A solution is termed *strong* if it satisfies $J(x(\cdot)) \leq J(y(\cdot))$ for all arcs $y(\cdot)$ satisfying only (1.3). Some of the fundamental classical necessary conditions for optimal solutions are the *Euler-Lagrange* equation, *Hamiltonian* (Canonical) conditions, and the *Weistrass* condition. In the following classical definitions we consider a piecewise smooth function $x : [a,b] \to \mathbb{R}^n$. The Euler-Lagrange equation in its classical form, for systems without time delay, states that assuming $x(\cdot)$ is a weak local minimum, the following equality holds

$$L_v(t, x(t), \dot{x}(t)) = c + \int_a^t L_x(s, x(s), \dot{x}(s))\, ds,$$

assuming L to be coercive and C^2 with $L_{vv} > 0$. Suppose that $x(\cdot)$ and $p(\cdot)$ are related by $p(t) = L_v(t, x(t), \dot{x}(t))^T$, where T stands for transpose, making $p(t)$ a column vector.

The *Hamiltonian* is a mapping $H : [a,b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by

$$H(t, x, p) := \sup_{v \in R^n} \big\{ \langle p, v \rangle - L(t, x, v) \big\}.$$

The classical Hamiltonian condition states that the Euler-Lagrange equation $\dot{p}(t) = L_x(t, x(t), \dot{x}(t))^T$ holds if and only if $x(\cdot)$ and $p(\cdot)$ satisfy the Hamiltonian system

$$\begin{aligned} \dot{x}(t) &= \nabla_p H(t, x(t), p(t))^T \\ -\dot{p}(t) &= \nabla_x H(t, x(t), p(t))^T. \end{aligned}$$

The Weistrass condition states that having x a strong local minima and v its derivative then $L(t, x, w) - L(t, x, v) - L_v(t, x, v)(w - v) \geq 0$ for all $w \in \mathbb{R}^n$.

1.4 Weak Invariance

An autonomous differential inclusion is of the form

$$\dot{x} \in F(x),$$

where F is a set valued map which associates with any point $x \in \mathbb{R}^n$ a set $F(x) \subset \mathbb{R}^n$. A system (S, F) consisting of a closed set S and a multifunction F mapping $\mathbb{R}^n$ to nonempty, compact, convex subsets of $\mathbb{R}^n$, is said to be *weakly invariant* provided that for all $x_0 \in S$, there is a trajectory x satisfying the differential inclusion $\dot{x}(t) \in F(x(t))$ for $t \in [0, \infty)$ so that

$$x(0) = x_0, \qquad x(t) \in S \quad \forall t \geq 0.$$

We will assume that F is upper semi-continuous and satisfies the following growth condition for some positive constants θ and c:

$$v \in F(x) \Rightarrow \|v\| \leq \theta \|x\| + c.$$

The *lower Hamiltonian* h corresponding to F is a function $h : I\!R \times I\!R^n \times I\!R^n \to I\!R$ defined by

$$h(t, x, p) = \min_{v \in F(t,x)} \langle p, v \rangle.$$

Chapter 2

The Problem of Bolza and Dynamical Systems with Time Delay

Before we begin our study of necessary conditions for the generalized problem of Bolza with time delay, we wish to establish its relationship to dynamical systems and control. The generalized Bolza problem without time delay contains a functional of the form

$$\bar{\Lambda}\big(x(\cdot)\big) := \ell\big(x(0), x(T)\big) \; + \; \int_0^T \; L\big(t, x(t), \dot{x}(t)\big)\, dt,$$

where $\ell : I\!R^n \times I\!R^n \to (-\infty, \infty]$ and $L : [0,T] \times I\!R^n \times I\!R^n \to (-\infty, \infty]$ are the given data and are allowed to take on $+\infty$ values as a means to incorporate constraints. The Bolza problem is

$$(P_B) \qquad\qquad \text{minimize } \bar{\Lambda}\big(x(\cdot)\big) \tag{2.1}$$

over $x(\cdot) \in \mathcal{AC}[0,T]$ (= the absolutely continuous arcs defined from $[0,T]$ into $I\!R^n$). This problem has been shown to encompass several other problems in dynamic optimization [2, 28]. In particular, it encompasses the general problem in the calculus of variations, (P_L), the differential inclusion problem, (P_D), and the measurable control problem, (P_C). We state these as follows:

(P_L) min $\int_0^T L(t, x(t), \dot{x}(t))\, dt$, for $x(0) = A$, $x(t) = B$, L locally Lipschitz

(P_D) min $\ell(x(0), x(T))$ for $\dot{x} \in F(t, x(t))$

(P_C) min $\int_0^T g(t, x(t), w(t))$ over measurable controls w taking values in some set $\mathcal{U}$. Here, $\dot{x}$ satisfies $\dot{x}(t) = f(t, x(t), w(t))$ *a.e.*, $x(0) = A$, $x(T) = B$.

Since (P_B) is the general case, one would like to derive necessary conditions for that case and obtain the others as Corollaries. However, until the work by Clarke in [2], necessary conditions for P_B had not been directly derived. The equivalence of P_B and P_C was established by Rockafellar [28]. We consider whether this equivalence continues to hold first, when only state time delay is present and second, when both state and velocity variables with time delay appear.

2.1 The State Time Delay Case: Equivalence of P_{DB} and P_{DC}

In this section, we consider the relationship between the delayed problem of Bolza, (P_{DB}), and the measurable control problem with time delay in the state variable, (P_{DC}). We assume that $\mathcal{X} \subset I\!R^n$, $\mathcal{U} \subset I\!R^m$ is a measurable set, and consider the measurable control functions $u(\cdot) : [0,T] \to \mathcal{U}$. In the state time-delay case, the generalized problem of Bolza P_{DB} and the delayed measurable control problem P_{DC} are defined as follows:

$$(P_{DB}) \qquad \min_{x(\cdot)\in\mathcal{X}} \int_0^T L(x(t), x(t-\Delta), \dot{x}(t))\, dt \tag{2.2}$$

$$(P_{DC}) \qquad \min_{x(\cdot)\in\mathcal{X}, u(\cdot)\in\mathcal{U}} \int_0^T L_0(x(t), x(t-\Delta), u(t))\, dt \tag{2.3}$$

$$\dot{x}(t) = f(x(t), x(t-\Delta), u(t))\, a.e., \tag{2.4}$$

where each $L : I\!R^n \times I\!R^n \times I\!R^n \to I\!R$ and $L_0 : I\!R^n \times I\!R^n \times I\!R^m \to I\!R$ is lower semicontinuous. We assume that $f : I\!R^n \times I\!R^n \times I\!R^m \to I\!R^n$ is Lipschitz and $u(\cdot) \in \mathcal{U}$ for almost all $t \in [0,T]$ (i.e. 'a.e.'). Also, $\Delta > 0$ is a constant and we define $x(t) = c(t)$ for each $t \leq 0$, where $c(\cdot) \in L^2[-\Delta, 0]$. Throughout, we define $v(\cdot) := \dot{x}(\cdot)$ as is standard in the calculus of variations and write $x(t-\Delta)$ as $x_\Delta(t)$. Unless otherwise noted, all equalities and inequalities are assumed to hold almost everywhere $t \in [0,T]$ (a.e.). In the following proposition, we show that these two problems are equivalent.

Proposition 2.1.1. *The delayed problem of Bolza P_{DB} is equivalent to the delayed measurable control problem P_{DC}*

Proof. We first show that given a solution $(\bar{x}(\cdot), \bar{u}(\cdot))$ to P_{DC}, $\bar{x}(\cdot)$ solves P_{DB}. Define

$$L(x, y, v) := \min_{u \in \mathcal{U}} \big\{ L_0(x, y, u) : v = f(x, y, u) \big\}. \tag{2.5}$$

Suppose that $(\bar{x}(\cdot), \bar{u}(\cdot))$ solves P_{DC}, then $L_0(\bar{x}(t), \bar{x}_\Delta(t), \bar{u}(t)) \leq$ $L_0(x(t), x_\Delta(t), u(t))$ for all $(x(\cdot), u(\cdot))$. By our definition (2.5) of $L(x, y, v)$, and letting $\bar{v}(\cdot) = \dot{\bar{x}}(\cdot)$ we have that for almost all t

$$\begin{aligned} &L(\bar{x}(t), \bar{x}_\Delta(t), \bar{v}(t)) = \\ &\quad \min_{u \in \mathcal{U}} \big\{ L_0(\bar{x}(t), \bar{x}_\Delta(t), u) : \bar{v}(t) = f(\bar{x}(t), \bar{x}_\Delta(t), u) \big\} = L_0(\bar{x}(t), \bar{x}_\Delta(t), \bar{u}(t)). \end{aligned} \tag{2.6}$$

The last equality follows since $(\bar{x}(\cdot), \bar{u}(\cdot))$ solves P_{DC}. We prove that $\bar{x}(\cdot)$ solves P_{DB}. Suppose not. Then there exists $\tilde{x}(\cdot) \in \mathcal{X}$ so that for almost all t

$$L(\tilde{x}(t), \tilde{x}_\Delta(t), \tilde{v}(t)) < L(\bar{x}(t), \bar{x}_\Delta(t), \bar{v}(t)). \tag{2.7}$$

From (2.5),

$$L(\tilde{x}(t), \tilde{x}_\Delta(t), \tilde{v}(t)) = \min_{u \in \mathcal{U}} \big\{ L_0(\tilde{x}(t), \tilde{x}_\Delta(t), u) : \tilde{v}(t) = f(\tilde{x}(t), \tilde{x}_\Delta(t), u) \big\}$$

and since $\mathcal{U}$ is measurable and $L_0(\cdot, \cdot, \cdot)$ is lower semicontinuous there exists $\tilde{u}_t \in \mathcal{U}$ that minimizes the set above. Using this and (2.6,2.7) above, we obtain for almost all t

$$\begin{aligned} L(\tilde{x}(t), \tilde{x}_\Delta(t), \tilde{u}_t) &= L(\tilde{x}(t), \tilde{x}_\Delta(t), \tilde{v}(t)) \\ &< L(\bar{x}(t), \bar{x}_\Delta(t), \bar{v}(t)) = L_0(\bar{x}(t), \bar{x}_\Delta(t), \bar{u}(t)). \end{aligned}$$

But, this contradicts our assumption that $(\bar{x}(\cdot), \bar{u}(\cdot))$ solves P_{DC} and thus we must have that $\bar{x}(\cdot)$ solves P_{DB}.

To show that given a solution $\bar{x}(\cdot)$ to P_{DB} we can find $\bar{u}(\cdot) \in \mathcal{U}$ so that $(\bar{x}(\cdot), \bar{u}(\cdot))$ solves P_{DC} we let $m = n$ and define $u(t) := \dot{x}(t) = f(x(t), x_\Delta(t), u(t))$. Now, consider the set

$$\mathcal{U}_t := \big\{ u(\cdot) \in \mathcal{U} : L(x(t), x_\Delta(t), v(t)) = L_0(x(t), x_\Delta(t), u(t)) \big\}.$$

Since the function $t \mapsto \mathcal{U}_t$ is measurable, we can obtain a measurable selection so that our set $\mathcal{U}_t$ is non-empty. We now consider our controls restricted to $\mathcal{U}_t$. Suppose that $\bar{x}(\cdot)$ solves P_{DB}, then $L(\bar{x}(t), \bar{x}_\Delta(t), \bar{v}(t)) \leq L(x(t), x_\Delta(t), v(t))$ for all $x(\cdot)$. Suppose that $(\tilde{x}(\cdot), \tilde{u}(\cdot))$ solves P_{DC}. Since, $\tilde{u}(\cdot) = \dot{\tilde{x}}(\cdot) = \tilde{v}(\cdot)$ and $\tilde{u}(\cdot) \in \mathcal{U}_t$, we then have that for almost all $t \in [0, T]$

$$\begin{aligned} L(\tilde{x}(t), \tilde{x}_\Delta(t), \tilde{v}(t)) &= L_0(\tilde{x}(t), \tilde{x}_\Delta(t), \tilde{u}(t)) \\ &\leq L_0(\bar{x}(t), \bar{x}_\Delta(t), \bar{u}(t)) = L(\bar{x}(t), \bar{x}_\Delta(t), \bar{v}(t)). \end{aligned}$$

But, since $\bar{x}(\cdot)$ solves P_{DB},

$$\begin{aligned} L_0(\bar{x}(t), \bar{x}_\Delta(t), \bar{u}(t)) &= L(\bar{x}(t), \bar{x}_\Delta(t), \bar{v}(t)) \\ &\leq L(\tilde{x}(t), \tilde{x}_\Delta(t), \tilde{v}(t)) = L_0(\tilde{x}(t), \tilde{x}_\Delta(t), \tilde{u}(t)), \end{aligned}$$

hence it must be that $L_0(\bar{x}(t), \bar{x}_\Delta(t), \bar{u}(t)) = L_0(\tilde{x}(t), \tilde{x}_\Delta(t), \tilde{u}(t))$. And thus, the minimum value for P_{DC} is attained with $(\bar{x}(\cdot), \bar{u}(\cdot)) \in \mathcal{X} \times \mathcal{U}_t$. This completes the proof of our proposition. $\square$

Although the argument in the above proposition uses a fixed constant time delay $\Delta > 0$, a similar argument may be used when $\Delta(\cdot)$ is a function of bounded variation. We can see that equivalence between the generalized problem of Bolza and the measurable control problem is clear in both the case without time delay [28] and, by the above proposition, in the case of time delay in the state variable.

However, in the neutral problem, where the Lagrangian L also depends on $\dot{x}(t-\Delta)$, this equivalence does not seem to hold. One can show that the problem of Bolza encompasses the measurable control problem but the converse is not true.

2.2 The Neutral Case

Consider the neutral problem of Bolza (P_{NB}) and the measurable control problem with delay in both the state and control variables (P_{NC}). They are defined as follows:

$$(P_{NB}) \qquad \min_{x(\cdot)\in\mathcal{X}} \int_0^T L(x(t), x(t-\Delta), \dot{x}(t), \dot{x}(t-\Delta))\, dt, \tag{2.8}$$

$$(P_{NC}) \qquad \min_{x(\cdot)\in\mathcal{X}} \int_0^T L_0(x(t), x(t-\Delta), u(t), u(t-\Delta))\, dt \tag{2.9}$$

$$\text{for} \qquad \dot{x}(t) = f(x(t), x(t-\Delta), u(t), u(t-\Delta))$$

where each $L : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to (-\infty, \infty]$ and $L_0 : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m \to (-\infty, \infty]$ are lower semicontinuous in all variables. The delay $\Delta > 0$ is a fixed constant. We assume that $f : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^n$ is Lipschitz, $u(t) \in \mathcal{U}$ for almost all $t \in [0, T]$ and $u(t) := 0$ for all $t < 0$. We define $x(t) := c(t)$ for almost all $t \in [-\Delta, 0]$, where $c(\cdot)$ is an absolutely continous function in $[\Delta, 0]$.

We proceed to show that P_{NB} encompasses P_{NC}. We again write $x_\Delta(t) := x(\cdot - \Delta)$ and $u_\Delta(t) := u(\cdot - \Delta)$.

Proposition 2.2.1. *If $x(\cdot)$ solves P_{NB} then there exists $u(\cdot) \in \mathcal{U}$ so that $(x(\cdot), u(\cdot))$ solves P_{NC}*

Proof. To show that given a solution $\bar{x}(\cdot)$ to P_{NB} we can find $\bar{u}(\cdot) \in \mathcal{U}$ so that $(\bar{x}(\cdot), \bar{u}(\cdot))$ solves P_{NC}, we let $m = n$ and define $u(t) := \dot{x}(t) = f(x(t), x_\Delta(t), u(t), u_\Delta(t))$. Consider the set

$$\mathcal{U}_t := \big\{u(\cdot) \in \mathcal{U} : L(x(t), x_\Delta(t), v(t), v_\Delta(t)) = L_0(x(t), x_\Delta(t), u(t), u_\Delta(t))\big\}.$$

Since the function $t \mapsto \mathcal{U}_t$ is measurable, we can obtain a measurable selection so that our set $\mathcal{U}_t$ is non-empty. We now consider our controls restricted to $\mathcal{U}_t$. Suppose that $\bar{x}(\cdot)$ solves P_{NB}, then $L(\bar{x}(t), \bar{x}_\Delta(t), \bar{v}(t), \bar{v}_\Delta(t)) \leq L(x(t), x_\Delta(t), v(t), v_\Delta(t))$ for all $x(\cdot)$. Suppose that $(\tilde{x}(\cdot), \tilde{u}(\cdot))$ solves P_{NC}, then

$$L_0(\tilde{x}(t), \tilde{x}_\Delta(t), \tilde{u}(t), \tilde{u}_\Delta(t)) \leq L_0(\bar{x}(t), \bar{x}_\Delta(t), \bar{u}(t), \bar{u}_\Delta(t)).$$

But, since $\bar{x}(\cdot)$ solves P_{NB} and each of $\tilde{u}$ and $\bar{u}$ are elements of $\mathcal{U}_t$, the following relationship holds almost everywhere,

$$\begin{aligned} L_0(\bar{x}(t), \bar{x}_\Delta(t), \bar{u}(t), \bar{u}_\Delta(t)) &= L(\bar{x}(t), \bar{x}_\Delta(t), \bar{v}(t), \bar{v}_\Delta(t)) \\ &\leq L(\tilde{x}(t), \tilde{x}_\Delta(t), \tilde{v}(t), \tilde{v}_\Delta(t)) = L_0(\tilde{x}(t), \tilde{x}_\Delta(t), \tilde{u}(t), \tilde{u}_\Delta(t)). \end{aligned}$$

Hence it must be that

$$L_0(\bar{x}(t), \bar{x}_\Delta(t), \bar{u}(t), \bar{u}_\Delta(t)) = L_0(\tilde{x}(t), \tilde{x}_\Delta(t), \tilde{u}(t), \tilde{u}_\Delta(t)).$$

Therefore, the minimum value for P_{NC} is attained with $(\bar{x}(\cdot), \bar{u}(\cdot)) \in \mathcal{X} \times \mathcal{U}_t$. This completes our proof. $\square$

If we were to attempt the same method of proof, as in Proposition 1, to show that a solution to P_{NC} also solves P_{NB} we would define

$$\begin{aligned} &L(x(t), x_\Delta(t), v(t), v_\Delta(t)) := \\ &\qquad \min_{u_1, u_2 \in \mathcal{U}} \big\{ L_0(x(t), x_\Delta(t), u_1, u_2) : v(t) = f(x(t), x_\Delta(t), u_1, u_2) \big\}. \end{aligned}$$

Although the above minimum exists, if it is attained at say $\bar{u}_1, \bar{u}_2$, there is no reason why $\bar{u}_2(t) = \bar{u}_1(t - \Delta)$. This makes it possible for $(\bar{x}(\cdot), \bar{u}(\cdot))$ to solve P_{NC} while $\bar{x}(\cdot)$ fails to be the solution to P_{NB}. We provide a simple linear example for which this occurs.

Let $u \in \mathcal{U}$ with $u(t) = t + b$ for some $b \in I\!R$. We will optimize over $x \in \mathcal{X}$. We let $\Delta = 1$ and $\dot{x}(t) = f(x(t), u(t), u(t-1)) = Ax(t) + B(u(t) - u(t-1))$, where A, B are positive real numbers and the initial condition is given by $x(0) = x_0$. In this case f has one solution $x(t) = \mathrm{e}^{tA}x_0 + B(\mathrm{e}^{tA} - 1)$. If we let

$$L_0(x(t), u(t), u(t-1)) = x^2(t) + (u(t) - u(t-1))^2 = x^2(t) + 1,$$

then the minimum in P_{NC} is attained at
$(\bar{x}(t), \bar{u}(t)) = (\mathrm{e}^{tA}x_0 + B(\mathrm{e}^{tA} - 1), t)$. Here any control $u(t) = t + b$ would have satisfied since L_0 is independent of $u(\cdot)$. Define

$$\begin{aligned} L(x(t), v(t), v_\Delta(t)) &= \min_{u_1, u_2 \in \mathcal{U}} \big\{ L_0(x(t), u_1, u_2) : v(t) = f(x(t), u_1, u_2) \big\} \\ &= \min_{u_1, u_2 \in \mathcal{U}} \big\{ x^2(t) + (u_1 - u_2)^2 : v(t) = f(x(t), u_1, u_2) \big\}, \end{aligned}$$

with initial condition $x(0) = x_0$ for $v(t)$. For any u_1, u_2 in $\mathcal{U}$, $\bar{x}^2(t) + (u_1 - u_2)^2 \geq \bar{x}^2(t)$. Minimizing over $u_1, u_2 \in \mathcal{U}$ we obtain

$$L(\bar{x}(t), \bar{v}(t), \bar{v}_\Delta(t)) \geq \bar{x}^2(t) \geq (e^{tA}x_0)^2 + (B(e^{tA} - 1))^2.$$

The last inequality holds since $\bar{x}(t) = \mathrm{e}^{tA}x_0 + B(\mathrm{e}^{tA} - 1)$. But, $v(t) = f(x(t), u_1, u_2)$ has a unique solution $\tilde{x}(t) = \mathrm{e}^{tA}x_0 + (u_2 - u_1)B(\mathrm{e}^{tA} - 1)$ so that

$$\begin{aligned} L(\tilde{x}(t), \tilde{v}(t), \tilde{v}_\Delta(t)) &= \min_{u_1, u_2 \in \mathcal{U}} \big\{ \tilde{x}^2(t) + (u_1 - u_2)^2 \big\} \leq (e^{tA}x_0)^2 \\ &< (e^{tA}x_0)^2 + (B(e^{tA} - 1))^2 < L(\bar{x}(t), \bar{v}(t), \bar{v}_\Delta(t)). \end{aligned}$$

Therefore, $\bar{x}(t)$ does not solve P_{NB}, where the first inequality holds because we can always take $u_1 = u_2$ when evaluating the minimum.

To be precise, our example simply shows that our method of proof does not work but it does not show the failure of equivalence for these neutral problems. The relationship between the two problems is not clear but we suspect that the two problems are not equivalent in general.

Chapter 3

Value Functions and a Decoupling Technique

3.1 Optimal Control and Value Functions

Consider the constrained optimization problem of minimizing a function $f(x)$ over all points x subject to $h(x) - \alpha = 0$, where $f : \mathbb{R}^n \mapsto \mathbb{R}$ and $h : \mathbb{R}^n \mapsto \mathbb{R}^m$ are given and are smooth. We refer to this problem as $P(\alpha)$. We associate with this problem the value function $V(\cdot)$ whose value $V(\alpha)$ is the infimum of $P(\alpha)$. In general, V may take values in $[-\infty, \infty]$, having value ∞ when there are no feasible values x satisfying the constraint $h(x) - \alpha = 0$. By definition we observe that for any $x \in \mathbb{R}^n$, $f(x) \geq V(h(x)) = V(\alpha)$ and equality holds for any x solving $P(\alpha)$.

Suppose x_0 solves $P(\alpha)$. Then, we have that $x \mapsto f(x) - V(h(x))$ attains a minimum at x_0. Thus, its derivative $f'(x_0) + \langle \nabla V(0), h'(x_0) \rangle = 0$. This is related to the Lagrange multiplier rule. However, the problem with this line of reasoning is that we are not guaranteed the existence of $\nabla V(0)$. Instead, we may use non-smooth analysis results to show that for a vector ζ in the proximal subgradient of V at 0, i.e. $\zeta \in \partial_P V(0)$, we can obtain that if x_0 solves $P(0)$, then $f'(x_0) + \langle \zeta, h'(x_0) \rangle = 0$ holds. Still, we must consider whether $\partial_P V(0)$ is non-empty. Under the following growth condition:

The set $\{x \in \mathbb{R}^n : f(x) \leq r, \|h(x)\| \leq s\}$ is bounded for every $r, s \in \mathbb{R}$, we may conclude that $V(\cdot)$ is l.s.c. and obtain our desired result as follows.

Here are some results obtained concerning the set $\partial_P V(0)$. Assuming the above growth condition and that $V(0) < \infty$, the following hold:

1. A solution to $P(0)$ exists. (This holds assuming only that $V(0) < \infty$).

2. There exists a sequence $\{\alpha_i\} \to 0$ with $V(\alpha_i) \to V(0)$ and points $\zeta_i \in \partial_P V(\alpha_i)$, x_i solving $P(\alpha_i)$, so that

$$f'(x_i) + \zeta^T h'(x_i) = 0$$

for each i. This result follows from the Density theorem.

3. Suppose that for each feasible x, the Jacobian $h'(x)$ is of maximal rank. Then $V(\cdot)$ is Lipschitz near zero, and we have

$$\emptyset \neq \partial_L V(0) \subset \bigcup \big\{\zeta \in I\!R^m : f'(x) + \zeta^T h'(x) = 0\big\},$$

where the union is taken over all feasible x.

4. If we assume normality of the solution x,
i.e. $0 \in \partial_L \langle \zeta, h(\cdot) \rangle)(x) \Rightarrow \zeta = 0$, then we get the same result as 3. for f, h locally Lipschitz.

Complete proofs showing the derivation of the above claims concerning $\partial_P V(0)$ may be found in ([3], pp. 103-108).

3.2 Clarke's Decoupling Principle

In our approach, we use value functions together with a decoupling principle introduced by Clarke [2] in order to derive necessary conditions for optimal solutions of constrained problems with time delay. In his paper *A decoupling principle in the calculus of variations* [2], Clarke considers the generalized problem of Bolza (2.1) and introduces these standards assumptions:

(C1) ℓ is lower semicontinuous and bounded below;

(C2) At least one of the following sets is bounded:

$\{\gamma : \ell(\gamma, \tau) < \infty \text{ for some } \tau\}$ $\qquad$ $\{\tau : \ell(\gamma, \tau) < \infty \text{ for some } \gamma\}$

(C3) $L(t, x, v)$ is lower semicontinuous in (x, v), is $\mathcal{L}\times\mathcal{B}$ measurable in t and (x, v), and is convex in v.

(C4) There exists a function $\theta : [0, \infty) \to I\!R$ satisfying $\lim_{r\to\infty} \theta(r)/r = \infty$ and so that

$$L\big(t, x, v\big) \geq \theta\big(|v|\big) \qquad \text{for all } v \in I\!R^n.$$

Using the "decoupling" principle Clarke obtains necessary conditions for optimal solutions of the Bolza problem. The term "decoupling" refers to a general optimization technique that can be roughly described as follows: additional free (i.e. nonconstrained) variables are introduced to parameterize a family of auxiliary problems and a proximal subgradient of the value function associated to these parameters will be the predecessor of a Lagrange multiplier of the original constrained problem. In our context, the new parameters can be subsequently used as perturbations to a given optimal arc and are related to the constraints of the original problem through the subgradient inequalities. Clarke's beautiful idea in [2], was to treat the condition that $\dot{x}(t)$ is the derivative of a trajectory $x(t)$ as a constraint. This leads to considering a minimization of a Bolza-type functional with the minimization taken over functions $\big(u(\cdot), v(\cdot)\big) \in L^2[0, T] \times L^1[0, T]$ under the constraint $u(t) = x_0 + \int_0^t v(s)\, ds$. Under additional assumptions, the subgradients with respect to these parameters can be shown to limit to adjoint arcs for the original problem as the parameters approach 0. The decoupling technique is also used as a tool in the derivation of necessary conditions in ([3],pp.230-243).

Chapter 4

Existence of Solutions to Neutral Bolza Problems

The existence of a solution for the general problem of Bolza without delay was established by Rockafellar [28] using the direct method in the Calculus of Variations. In what follows, we use a similar method of proof to obtain existence of solutions for our neutral problem of Bolza. The generalized problem of Bolza with time delay in the state and velocity variables involves an integral functional of the form

$$\Lambda\big(x(\cdot)\big) := \int_0^T L\big(t, x(t), x(t-\Delta), \dot{x}(t), \dot{x}(t-\Delta)\big)\, dt, \tag{4.1}$$

where $L : [0,T] \times I\!R^n \times I\!R^n \times I\!R^n \times I\!R^n \to (-\infty, \infty]$ is given data and, as is now standard in variational analysis [2, 28], is allowed to take on $+\infty$ values as a means to incorporate constraints. The infinite penalization technique has an obvious and natural extension to delay problems. The time delay $\Delta > 0$ is a fixed constant. The delayed velocity variable $\dot{x}(t)$ is set equal to 0 in the interval $[-\Delta, 0]$, and the given fixed initial state variable (the "initial tail") $c(\cdot) : [-\Delta, 0] \to I\!R^n$ is assumed to be in $L^2[-\Delta, 0]$. Notice that the definition of $\dot{x}(s) = 0$ for each $s < 0$ is an arbitrary choice of a "tail "for the derivative. This "tail"may be replaced by any piecewise continuous function $\xi(\cdot)$. Another option is to define $\dot{x}(s) = \dot{c}(s)$ for all $s \in [-\Delta, 0]$ but this would require that $c(\cdot)$ be absolutely continuous. The generalized Bolza problem with time delay is the following optimization problem:

$$\text{minimize } \ell\big(x(T)\big) \; + \; \Lambda\big(x(\cdot)\big) \tag{4.2}$$

over $x(\cdot) \in \mathcal{AC}[0,T]$ (= the absolutely continuous arcs defined from $[0,T]$ into $I\!R^n$), and where $\ell : I\!R^n \to (-\infty, \infty]$ and $x(t) = c(t)$ for all $t \in [-\Delta, 0]$. The usual conventions of extended-valued real arithmetic are in use here; in particular, $\infty - \infty = \infty$. An arc $x(\cdot) \in \mathcal{A}C[0,T]$ is said to be *feasible* for (4.2) provided the sum $\ell\big(x(T)\big) + \Lambda\big(x(\cdot)\big)$ is finite.

4.1 Assumptions

In order to obtain existence of solutions for the Neutral Bolza Problem, we assume the following basic assumptions.

(H1) ℓ is lower semi-continuous and bounded below;

(H2) $L(t,x,y,v,w)$ is lower semi-continuous in (x,y,v,w), is $\mathcal{L}\times\mathcal{B}$-measurable on $[0,T]\times I\!R^{4n}$, and is (jointly) convex in (v,w).

(H3) There exists a nondecreasing function $\theta : [0,\infty) \to I\!R$ satisfying $\lim_{r\to\infty} \theta(r)/r = \infty$ so that

$$L\big(t,x,y,v,w\big) \geq \theta\big(|v|\big) + \theta\big(|w|\big) \qquad \text{for all } v,w \in I\!R^n.$$

These assumptions are the natural extension of those used by Clarke [2] and in [28] and [30], adding delays in [28] and ∞ in [30].

4.2 The Dunford-Pettis Criterion

For the sake of completeness, we provide a proof of the Dunford-Pettis Criterion, which is used in our existence proof. The proof is adapted from *Functional Analysis Theory and Applications* by Edwards [8]. Recall that the weak topology on a Banach space E, with dual E' is the weakest topology on E relative to which each of the linear forms $x \to \langle x, x'\rangle$, $x \in E$, $x' \in E'$, are continuous. The weak* topology is the weakest topology on E' relative to which each of the linear forms $x \to \langle x', x''\rangle$, $x'' \in E''$, are continuous.

Theorem 4.2.1. *Let T be a compact set with positive Radon measure μ. Let P be a subset of L^1. Suppose*

(1) $\sup\{\int |f|\, d\mu : f \in P\} < \infty$

(2) Given $\epsilon > 0$, there exists $\delta > 0$ such that if $A \subset T$,$\mu(A) < \delta$ then $\sup\{\int_A |f|\, d\mu : f \in P\} \leq \epsilon$.

Then, P is weakly relatively compact.

Proof. By (1) P is bounded in L^1 and therefore it is bounded in the dual of L^∞. But, the bounded sets in the dual are exactly those sets that are weak* relatively compact. Therefore P is weak* relatively compact in the dual of $L^\infty, (L^\infty)^*$. Let $\bar{P}$ be the (compact) weak* closure of P in $(L^\infty)^*$. We must show that $\bar{P} \subset \mathcal{L}^1$ (or more precisely that each element in $\bar{P}$ can be identified with an element in L^1).

Let L be a continuous linear functional in $\bar{P}$. From (2) it follows that for any $\epsilon > 0$, there is a $\delta > 0$ so that for A as in (2), $g \in L^\infty$, $|g| \leq \mathcal{X}_A$, and for each $f \in P$

$$|\int fg\, d\mu| \leq \int_A |f| d\mu \leq \epsilon. \tag{4.3}$$

Since $L \in \bar{P}$, $|L(g)| \leq \epsilon$ when g is as above. Now, consider the restriction of L to the set of continuous functions with compact support in T, and denote this set as $\mathcal{K}(T)$. By the Riez representation theorem there exists a Radon measure λ so that $L(g) = \int g\, d\lambda$ when $g \in \mathcal{K}(T)$. Also, $\left|\int g\, d\lambda\right| = \left|L(g)\right| \leq \epsilon$ when $g \in \mathcal{K}(T), |g| \leq \mathcal{X}_A$, and $\mu(A) \leq \delta$. But, this means that $|\lambda|(A) \leq \sup |\int g d\lambda| \leq \epsilon$. Therefore, λ is absolutely continuous with respect to μ. By the Radon-Nykodym theorem there exists an integrable function f_0 so that $\lambda = f_0 \mu$. Hence,

$$L(g) = \int f_0 g\, d\mu \tag{4.4}$$

when $g \in \mathcal{K}(T)$.

We now wish to extend this characterization of L for all functions $g \in L^\infty$. In order to show this, choose a function $g \in L^\infty$. By Lusin's theorem we may find a function g_n so that $g_n = g$ for all t except in a set of small measure and with $\|g_n\|_\infty \leq \|g\|_\infty$. Choose a sequence of such functions $\{g_n\}$ and choose δ' so that for some N and $n > N$, when $\mu(\{t : g_n \neq g\}) < \delta'$, $\int_{\{t:g_n \neq g\}} |h| \, d\mu < \epsilon / \big(2\|g\|_\infty\big)$ for each h. Our desired result follows from the following claims:

Claim 4.2.1. $\int f_0 g_n \, d\mu \to \int f_0 g \, d\mu$

Proof. This can be observed in the following estimate.

$$\begin{aligned} \Big|\int f_0 g_n \, d\mu - \int f_0 g \, d\mu\Big| &\leq \int |f_0||g_n - g| \, d\mu \\ &= \int_{\{t:g_n \neq g\}} |f_0||g_n - g| \, d\mu < \epsilon \end{aligned} \tag{4.5}$$

where the last inequality follows from our choice of δ' and $f_0 \in L^1$. □

Claim 4.2.2. $L(g_n) \to L(g)$

Proof. Since L is a linear functional we have that $|L(g_n) - L(g)| = |L(g_n - g)|$. Now, since $g_n = g$ except in a set of measure δ' and for each $f \in P$,

$$\Big|\int f(g_n - g) \, d\mu\Big| \leq 2\|g\|_\infty \int_{\{t:g_n \neq g\}} |f| \, d\mu < \epsilon.$$

The last inequality follows from our choice of δ' and $f \in P$. But, since $L \in \bar{P}$, this means that under all these same properties $\big|L(g_n - g)\big| < \epsilon$ which proves our claim. □

The above claims combined imply that $L(g) = \int f_0 g \, d\mu$ for each $g \in L^\infty$. Thus, for each $L \in \bar{P}$ there exists $f_0 \in \mathcal{L}^1$ such that $L(g) = \int f_0 g \, d\mu$. Therefore, there is an injection $L^1 \hookrightarrow (L^\infty)^*$, which concludes our proof. □

4.3 An Existence Theorem

We now proceed to show existence of solutions for the neutral problem of Bolza (4.2).

Theorem 4.3.1. *If there exists at least one $x(\cdot) \in \mathcal{AC}[0,T]$ that is feasible for (4.2), then there exists an arc $\bar{x}(\cdot) \in AC[0,T]$ that solves (4.2).*

There are two main ingredients to the proof, and these are separately given in the following lemmas.

Lemma 4.3.2. *Suppose $\mathcal{X} \subseteq AC[0,T]$ and $\mathcal{V} \subseteq L^1[0,T]$ are nonempty so that*

$$\sup_{x(\cdot)\in\mathcal{X}, v(\cdot)\in\mathcal{V}} \int_0^T L(t, x(t), x(t-\Delta), v(t), v(t-\Delta))\,dt < K < \infty. \tag{4.6}$$

Then $\mathcal{V}$ is weakly sequentially precompact.

Proof. The crucial information is contained in the estimate (4.7). For every $v(\cdot) \in \mathcal{V}$, we have by (H3) and (4.6) that

$$\int_0^T \theta(|v(t)|) + \theta(|v(t-\Delta)|)\,dt < K.$$

Also, $\int_0^T \theta(|v(t)|)\,dt \le K - T\theta(0)$ since $\theta(0)$ is the min value of θ. Let $I \subseteq [0,T]$ be any measurable set and $R > 0$ so that $\theta(R) > 0$, and define $A := I \cap \{t : |v(t)| \le R\}$ and $B := I \cap \{t : |v(t)| > R\}$. Then

$$\begin{aligned}
\int_I |v(t)|\,dt &= \int_A |v(t)|\,dt + \int_B |v(t)|\,dt \\
&\le Rm(I) + \int_B \frac{|v(t)|}{\theta(|v(t)|)}\theta(|v(t)|)\,dt \\
&\le Rm(I) + \int_B \sup_{r\ge R} \frac{r}{\theta(r)}\theta(|v(t)|)\,dt \\
&\le Rm(I) + (K - T\theta(0)) \sup_{r\ge R} \frac{r}{\theta(r)}.
\end{aligned}$$

Notice that $\theta(|v(t)|) \geq \theta(R) > 0$ for $t \in B$, so that dividing by $\theta(|v(t)|)$ is permitted. To summarize, there exists a constant k so that

$$\sup_{v(\cdot)\in\mathcal{V}} \int_I |v(t)\, dt \leq Rm(I) + k \sup_{r\geq R} \frac{r}{\theta(r)} \tag{4.7}$$

for all measurable $I \subseteq [0,T]$ and large R.

Now recall the Dunford-Pettis criterion (page 274 of [8]), which says that $\mathcal{V} \subseteq L^1[0,T]$ is weakly sequentially precompact if and only if

(i) $\sup_{v(\cdot)\in\mathcal{V}} \|v(\cdot)\|_1 < \infty$, and

(ii) For all $\varepsilon > 0$ there exists a $\delta > 0$ such that $m(I) < \delta$ implies $\int_I |v(s)|\, ds < \varepsilon$ for all $v(\cdot) \in \mathcal{V}$.

Since $\frac{r}{\theta(r)} \to 0$ as $r \to \infty$, (i) follows immediately from (4.7) by letting $I = [0,T]$. To see (ii), let $\varepsilon > 0$. There exists $R > 0$ such that $\sup_{r\geq R} \frac{r}{\theta(r)} < \frac{\varepsilon}{2k}$. Let $\delta = \frac{\varepsilon}{2R}$. If $m(I) < \delta$, then it follows from (4.7) that

$$\int_I |v(s)|\, ds \leq R\delta + k\frac{\varepsilon}{2k} < \varepsilon.$$

This shows (ii) holds and the lemma is proved. □

The second ingredient in the proof of Theorem 4.3.1 is the weak lower semicontinuity of the $\Lambda(\cdot)$ in (4.2), and for this purpose, we introduce the following natural modification of the maximized Hamiltonian:

$$H : I\!R \times I\!R^{4n} \to I\!R$$

$$H\big(t,x,y,p,q\big) := \sup_{(v,w)\in R^{2n}} \big\{\langle p,v\rangle + \langle q,w\rangle - L(t,x,y,v,w)\big\}. \tag{4.8}$$

As in nondelay problems ([28]), H is upper semi-continuous in (x,y,v,w), $\mathcal{L}\times\mathcal{B}$-measurable on $[0,T]\times I\!R^{4n}$, and is (jointly) convex in (p,q). The joint convexity of L in (v,w) implies the conjugacy relationship

$$L(t,x,y,v,w) = \sup_{(p,q)\in\mathbb{R}^{2n}} \big\{\langle p,v\rangle + \langle q,w\rangle - H(t,x,y,p,q)\big\}.$$

The following proposition says, in essence, that the conjugacy "goes through" an integral, and a proof can be found in Theorem 1 of [27].

Proposition 4.3.3. *Suppose $x(\cdot)$, $y(\cdot)$, $v(\cdot)$, and $w(\cdot)$ are measurable satisfying $L\big(x(\cdot),y(\cdot),v(\cdot),w(\cdot)\big) \in L^1[0,T]$. Then $\int_0^T L(t,x(t),y(t),v(t),w(t))\,dt$ is equal to the supremum of*

$$\int_0^T \big[\langle p(t),v(t)\rangle + \langle q(t),w(t)\rangle - H(t,x(t),y(t),p(t),q(t))\big]\,dt$$

taken over $\big(p(\cdot),q(\cdot)\big)$ in L^∞ (= the bounded measurable functions defined from $[0,T]$ into $\mathbb{R}^{2n}$).

We proceed to show the lower semicontinuity of $\Lambda(\cdot)$. The notation $v_i(t) \xrightarrow{w} \bar{v}(t)$ means that $\{v_i(\cdot)\}$ weakly converges to $\bar{v}(\cdot)$ in $L^1[0,T]$.

Lemma 4.3.4. *Suppose sequences $\{x_i(\cdot)\} \subseteq L^2[0,T]$ and $\{v_i(\cdot)\} \subseteq L^1[0,T]$ are such that $x_i(t) \to \bar{x}(t)$ for almost all t in $[0,T]$ and $v_i(t) \xrightarrow{w} \bar{v}(t)$. Then,*

$$\begin{aligned}\int_0^T L\big(t,\bar{x}(t),\bar{x}(t-\Delta),\bar{v}(t),\bar{v}(t-\Delta)\big)\,dt \\ \le \liminf_{i\to\infty} \int_0^T L\big(t,x_i(t),x_i(t-\Delta),v_i(t),v_i(t-\Delta)\big)\,dt.\end{aligned}$$

Proof. It follows from Proposition 4.3.3 that

$$
\begin{aligned}
&\liminf_{i\to\infty} \int_0^T L(t, x_i(t), x_i(t-\Delta), v_i(t), v_i(t-\Delta))\, dt \\
&\quad = \liminf_{i\to\infty} \sup_{(p(\cdot),q(\cdot))\in L^\infty} \left\{ \int_0^T \langle p(t), v_i(t)\rangle + \langle q(t), v_i(t-\Delta)\rangle \right. \\
&\qquad\qquad\qquad \left. - H\left(t, x_i(t), x_i(t-\Delta), p(t), q(t)\right)\, dt \right\} \\
&\quad \geq \sup_{(p(\cdot),q(\cdot))\in L^\infty} \liminf_{i\to\infty} \left\{ \int_0^T \langle p(t), v_i(t)\rangle + \langle q(t), v_i(t-\Delta)\rangle \right. \\
&\qquad\qquad\qquad \left. - H\left(t, x_i(t), x_i(t-\Delta), p(t), q(t)\right)\, dt \right\} \\
&\quad \geq \sup_{(p(\cdot),q(\cdot))\in L^\infty} \left\{ \int_0^T \langle p(t), \bar{v}(t)\rangle + \langle q(t), \bar{v}(t-\Delta)\rangle \right. \\
&\qquad\qquad\qquad \left. - \limsup_{i\to\infty} H\left(t, x_i(t), x_i(t-\Delta), p(t), q(t)\right)\, dt \right\} \\
&\quad \geq \sup_{(p(\cdot),q(\cdot))\in L^\infty} \left\{ \int_0^T \langle p(t), \bar{v}(t)\rangle + \langle q(t), \bar{v}(t-\Delta)\rangle \right. \\
&\qquad\qquad\qquad \left. - H\big(t, \bar{x}(t), \bar{x}(t-\Delta), p(t), q(t)\big)\, dt \right\} \\
&\quad = \int_0^T L\left(t, \bar{x}(t), \bar{x}(t-\Delta), \bar{v}(t), \bar{v}(t-\Delta)\right)\, dt.
\end{aligned}
$$

The second inequality is justified by Fatou's Lemma and $v_i(t) \xrightarrow{w} \bar{v}(t)$, the last inequality since H is upper semicontinuous, and the final equality by Proposition 4.3.3. □

Proof. (*of Theorem 4.3.1*) From our assumption of existence of a feasible $x(\cdot)$, we can select a minimizing sequence $\{x_i(\cdot)\} \subset \mathcal{AC}[0,T]$ so that for each i, $\ell(x_i(T)) + \Lambda\big(x_i(\cdot)\big) \leq \ell(x(T)) + \Lambda\big(x(\cdot)\big) < \infty$. Furthermore, $\ell(\cdot)$ is bounded below and $\ell(x_i(T)) + \Lambda\big(x_i(\cdot)\big)$ is bounded above so that $\Lambda\big(x_i(\cdot)\big)$ is bounded above, and so by Lemma 4.3.2 there exists $\bar{v}(\cdot) \in L^1[0,T]$ and a subsequence (which we do not relabel) satisfying $\dot{x}_i(\cdot) \xrightarrow{w} \bar{v}(\cdot)$. Defining $\bar{x}(\cdot) \in \mathcal{AC}[0,T]$ by

$$\bar{x}(t) = c(0) + \int_0^t \bar{v}(s)\, ds$$

(and set equal to $c(t)$ for $t \in [-\Delta, 0]$), we have $x_i(t) \to \bar{x}(t)$ for all $t \in [0, T]$. It follows from Lemma 4.3.4 and (H1) that

$$\ell\big(\bar{x}(T)\big) + \Lambda\big(\bar{x}(\cdot)\big) \leq \liminf_{i \to \infty} \ell\big(x_i(T)\big) + \Lambda\big(x_i(\cdot)\big)$$

Since $\{x_i(\cdot)\}$ is a minimizing sequence, it follows that $\bar{x}(\cdot)$ solves (4.2). $\square$

Chapter 5
The Decoupling Principle and Delay

We now consider the neutral problem of Bolza with a delay function of bounded variation appearing in the state and velocity variables of the Lagrangian L. We consider the following functional:

$$\Lambda\big(x(\cdot)\big) := \ell\big(x(T)\big) \ + \int_0^T L\big(t, x(t), x(t-\Delta(t)), \dot{x}(t), \dot{x}(t-\Delta(t))\big)\, dt, \tag{5.1}$$

where $\ell : I\!\!R^n \to (-\infty, \infty]$ and $L : [0,T] \times I\!\!R^n \times I\!\!R^n \times I\!\!R^n \times I\!\!R^n \to (-\infty, \infty]$ are the given data and are allowed to take on $+\infty$ values as a means to incorporate constraints. The delay is a function of bounded variation $\Delta : [0,T] \to [0, \Delta_0]$ with $|\dot{\Delta}(t)| \le K_\Delta$, $\Delta(0) = 0$, and where $\Delta_0 > 0$ and $K_\Delta > 0$ are fixed constants. We are also given the fixed initial data (the "initial tail") $c : [-\Delta_0, 0] \to I\!\!R^n$ that is assumed to be in $L^2[-\Delta_0, 0]$.

NPB with varying time delay is the following optimization problem:

$$\text{minimize } \Lambda\big(x(\cdot)\big) \tag{5.2}$$

over $x(\cdot) \in \mathcal{AC}[0,T]$, where $x(t) = c(t)$ and $\dot{x}(t) = 0$ for almost all $t \in [-\Delta_0, 0]$.

5.1 Decoupling and the Neutral Bolza Problem

In this section we show that Clarke's decoupling principle [2] can be extended to problem (5.2) by treating the delayed arcs $x(t-\Delta(t))$ and $\dot{x}(t-\Delta(t))$ as additional constraints, in which, $x(t-\Delta(t))$ and $\dot{x}(t-\Delta(t))$ are the time delays of the arcs $x(\cdot)$ and $\dot{x}(\cdot)$ respectively.

Let $\mathcal{X} := I\!\!R^n \times L^2[0,T] \times L^2[0,T] \times L^1[0,T] \times L^1[0,T]$. A new Bolza-type functional $\Gamma : \mathcal{X} \to (-\infty, \infty]$ that is similar to (5.1) is defined by

$$\Gamma\Big(\gamma, u(\cdot), w(\cdot), v_1(\cdot), v_2(\cdot)\Big) := \ell\big(\gamma\big) \ + \int_0^T L\big(t, u(t), w(t), v_1(t), v_2(t)\big)\, dt. \tag{5.3}$$

It is not difficult to show that problem (5.2) is equivalent to minimizing Γ over $\mathcal{X}$ subject to the constraints

$$\begin{aligned} u(t) &= x(0) + \int_0^t v_1(s)\,ds & (5.4)\\ \gamma &= x(0) + \int_0^T v_1(t)\,dt & (5.5)\\ v_2(t) &= v_1(t-\Delta(t)) & (5.6)\\ w(t) &= \xi(t)[c(t-\Delta(t)) - x(0)] + x(0) + \int_0^t v_2(s)[1-\dot{\Delta}(s)]\,ds, & (5.7) \end{aligned}$$

$\forall t \in [0,T]$. We define $\xi(t)$ as follows:

$$\xi(t) = \begin{cases} 1, & \text{if } t - \Delta(t) < 0; \\ 0 & \text{otherwise,} \end{cases} \tag{5.8}$$

and we define $v_1(t) = 0$ for each $t \in [-\Delta_0, 0]$. Indeed, (5.4) says that $x(t) := u(t)$ is absolutely continuous, (5.5) implies that γ is the endpoint $x(T)$, and (5.6)-(5.7) together say that $w(t) = x(t - \Delta(t))$ for $t \in [0,T]$. The following function $\mathcal{D} : \mathcal{X} \to I\!R$ is used to monitor how far an element $\big(\gamma, u(\cdot), w(\cdot), v_1(\cdot), v_2(\cdot)\big) \in \mathcal{X}$ is from satisfying (5.4)-(5.7):

$$\begin{aligned} &\mathcal{D}\big(\gamma, u(\cdot), w(\cdot), v_1(\cdot), v_2(\cdot)\big) = \\ &\quad \left|\gamma - x(0) - \int_0^T v_1(t)\,dt\right| + \int_0^T \left|u(t) - x(0) - \int_0^t v_1(s)\,ds\right|^2 dt \\ &+ \int_0^T \left|w(t) - \xi(t)[c(t-\Delta(t)) - x(0)] - x(0) - \int_0^t v_2(s)[1-\dot{\Delta}(s)]\,ds\right|^2 dt \\ &+ \int_0^T |v_2(t) - v_1(t-\Delta(t))|\,dt. \end{aligned}$$

It is clear that (5.4)-(5.7) hold if and only if $\mathcal{D}\big(\gamma, u(\cdot), w(\cdot), v_1(\cdot), v_2(\cdot)\big) = 0$. The decoupling principle is contained in the following theorem.

Theorem 5.1.1. *Suppose (H1)-(H3) from section 4.3 hold and $\epsilon > 0$. Then there exist a constant $\sigma > 0$, an element $\big(\bar{\gamma}, \bar{u}(\cdot), \bar{w}(\cdot), \bar{v}_1(\cdot), \bar{v}_2(\cdot)\big) \in \mathcal{X}$, and absolutely continuous arcs $p(\cdot)$ and $q(\cdot)$ defined on $[0,T]$ so that* *(a)* $\mathcal{D}\big(\bar\gamma, \bar u(\cdot), \bar w(\cdot), \bar v_1(\cdot), \bar v_2(\cdot)\big) < \epsilon$*;*

(b) $\Gamma\big(\bar\gamma, \bar u(\cdot), \bar w(\cdot), \bar v_1(\cdot), \bar v_2(\cdot)\big)$ *is within* ϵ *of the minimum value in (5.2);*

(c) for almost all $t \in [0, T]$*, the map*

$$\begin{aligned}(u, w, v_1, v_2) \mapsto L(t, u, w, v_1, v_2) &- \langle \dot p(t), u\rangle - \langle \dot q(t), w\rangle - \langle p(t), v_1\rangle \\ &- \langle q(t)[1 - \dot\Delta(t)], v_2\rangle \\ &+ \sigma\Big\{|u - \bar u(t)|^2 + |w - \bar w(t)|^2 + |v_1 - \bar v_1(t)|^2 + |v_2 - \bar v_2(t)|^2\Big\}\end{aligned}$$

is minimized at $(u, w, v_1, v_2) = \big(\bar u(t), \bar w(t), \bar v_1(t), \bar v_2(t)\big)$*; and*

(d) the map

$$\gamma \mapsto \ell(\gamma) + \langle \gamma, p(T)\rangle + \sigma|\gamma - \bar\gamma|^2$$

is minimized at $\gamma = \bar\gamma$*.*

Proof. Denote $P_{\eta,\alpha,\beta}$ as the problem of minimizing over all arcs $x(\cdot)$ the functional $\Lambda_{\eta,\alpha(\cdot),\beta(\cdot)}\big(x(\cdot)\big) :=$

$$\ell\big(x(T) + \eta\big) \; + \; \int_0^T L\big(t, x(t) + \alpha(t), x(t - \Delta(t)) + \beta(t), \dot x(t), \dot x(t - \Delta(t))\big)\, dt, \quad (5.9)$$

where $x(t)$ is set equal to $c(t)$ for $t \in [-\Delta_0, 0], \big(\eta, \alpha(\cdot), \beta(\cdot)\big) \in I\!R^n \times L^2[0,T] \times L^2[0,T]$, and where $\dot x(t-\Delta(t))$ and $\beta(\cdot) \in L^2[0,T]$ are each set equal to 0 whenever $t - \Delta(t) < 0$ to ensure we do not perturb the tail. Note that the integral in (5.9) is well-defined since it is a normal integrand [25].

Define a value function $V : I\!R^n \times L^2[0,T] \times L^2[0,T] \to (-\infty, \infty]$ by setting $V\big(\eta, \alpha(\cdot), \beta(\cdot)\big)$ as the minimum value in (5.9). Here $V\big(\eta, \alpha(\cdot), \beta(\cdot)\big) = \infty$ if there are no feasible arcs for $P_{\eta,\alpha,\beta}$. We begin with a lemma.

Lemma 5.1.2. *V is lower semicontinuous. If* $V\big(\eta, \alpha(\cdot), \beta(\cdot)\big) < \infty$ *then a solution to* $P_{\eta,\alpha,\beta}$ *exists.*

Proof. Choose $(\eta, \alpha(\cdot), \beta(\cdot)) \in I\!\!R^n \times L^2[0,T] \times L^2[0,T]$. To simplify notation, we define $\kappa := (\eta, \alpha(\cdot), \beta(\cdot))$. To proove that V is lower semicontinuous, i.e. $\liminf_{(\kappa_i \to \kappa)} V(\kappa_i) \geq V(\kappa)$, we consider two cases.

Case1:$V(\kappa) = \infty$

Choose a sequence κ_i converging almost everywhere, in t, to κ and suppose there is a subsequence, we do not relable, $\{\kappa_i\}$ so that $V(\kappa_i) < \infty$ for each i. Then, for each i, there exists a feasible arc x_i so that $\Lambda_{\kappa_i}(x_i) < \infty$. Choose a minimizing subsequence, without relabling, $\{x_i\}$ converging a.e. to $\bar{x}$. Consider the set of feasible arcs

$$\mathcal{B}_1 := \{x : \Lambda_{\kappa_i}(x) \leq \Lambda_{\kappa_i}(\bar{x})\}.$$

By lower semicotinuity of $\Lambda(\cdot)$ we know that, for each i, $\Lambda_{\kappa_i}(x_i) \geq \Lambda_{\kappa_i}(\bar{x})$. This means that $\Lambda_{\kappa_i}(\bar{x}) < \infty$. Taking the limit of Λ_{κ_i} and by the lower semicontinuity of $\Lambda(\cdot)$ this implies that $\Lambda_{\kappa}(\bar{x}) < \infty$. But this is a contradiction since $\infty = V(\kappa) \leq \Lambda_{\kappa}(\bar{x}) < \infty$. Thus it must be that for our original sequence $V(\kappa_i) = \infty$ for each i so that in this case V is lower semicontinuous.

Case2:$V(\kappa) < \infty$

Choose a sequence $\kappa_i \to \kappa$ a.e. and suppose that, for each i,
$V(\kappa_i) \leq V(\kappa) - \frac{\epsilon}{2}$. Then $V(\kappa_i) < \infty$ so, for each i, there exists a feasible arc x_i so that $\Lambda_{\kappa_i}(x_i) < \infty$. Thus, P_{κ_i} has a solution, say $\bar{x}_i$. From our asumption $\Lambda_{\kappa_i}(\bar{x}_i) \leq V(\kappa) - \frac{\epsilon}{2}$. Consider the set of feasible arcs

$$\mathcal{M}_2 := \{x : \Lambda_{\kappa_i}(x) \leq V(\kappa); \|\kappa_i - \kappa\| < \epsilon\}.$$

Choose a sequence $\{x_j\} \subset \mathcal{B}_2$ with $x_j \to \bar{x}$. Since $\mathcal{B}_2$ is closed $\bar{x} \in \mathcal{B}_2$. But then, $\lim \Lambda_{\kappa_i}(\bar{x}) = \Lambda_{\kappa}(\bar{x}) \leq V(\kappa)$. This implies that $\Lambda_{\kappa_i}(\bar{x}) \leq V(\kappa) - \frac{\epsilon}{2} \leq \Lambda_{\kappa}(\bar{x}) - \frac{\epsilon}{2}$. But this contradicts the lower semicontinuity of Λ. Hence, $V(\kappa_i) > V(\kappa)$ so that V is lower semicontinuous.

The existence of a solution to $P_{\eta,\alpha,\beta}$ follows from the Existence Theorem in Chapter 4. $\square$

The Density Theorem ([3], Theorem 3.1) of proximal analysis guarantees that for any $\epsilon > 0$ there exists $\big(\bar{\eta}, \bar{\alpha}(\cdot), \bar{\beta}(\cdot)\big) \in I\!R^n \times L^2[0,T] \times L^2[0,T]$ with $|\bar{\eta}| + \|\bar{\alpha}\|_2 + \|\bar{\beta}\|_2 < \epsilon$ and with $|V\big(\bar{\eta}, \bar{\alpha}(\cdot), \bar{\beta}(\cdot)\big) - V(0,0,0)| < \epsilon$. And that there exists $\big(\zeta, \phi(\cdot), \psi(\cdot)\big) \in I\!R^n \times L^2[0,T] \times L^2[0,T]$ with $\big(\zeta, \phi(\cdot), \psi(\cdot)\big) \in \partial_P V\big(\bar{\eta}, \bar{\alpha}(\cdot), \bar{\beta}(\cdot)\big)$, where $\partial_P V$ denotes the proximal subgradient of V. Notice that $\psi(t) = 0$ for all $t \in [-\Delta_0, 0]$ satisfying $t - \Delta(t) < 0$. By the proximal subgradient inequality we have that there exist $\delta' > 0$ and $\sigma' > 0$ such that for all $(\eta, \alpha, \beta) \in B((\bar{\eta}, \bar{\alpha}, \bar{\beta}), \delta')$ the following inequality holds.

$$\begin{aligned} V(\eta,\alpha,\beta) - V(\bar{\eta},\bar{\alpha},\bar{\beta}) + \sigma'\{|\eta-\bar{\eta}|^2 + \|\alpha-\bar{\alpha}\|_2^2 + \|\beta-\bar{\beta}\|_2^2\} \\ \geq \langle \zeta, \eta-\bar{\eta}\rangle + \langle \phi, \alpha-\bar{\alpha}\rangle + \langle \psi, \beta-\bar{\beta}\rangle \quad (5.10) \end{aligned}$$

Let $\bar{x}(\cdot)$ be an optimal solution to (5.9) with $\big(\eta, \alpha(\cdot), \beta(\cdot)\big) = \big(\bar{\eta}, \bar{\alpha}(\cdot), \bar{\beta}(\cdot)\big)$, which implies that $V(\bar{\eta},\bar{\alpha},\bar{\beta}) = \Lambda_{\bar{\eta},\bar{\alpha},\bar{\beta}}(\bar{x})$. It is given, from the definition of V, that for any arc $x(\cdot)$ one has $V(\eta,\alpha,\beta) \leq \Lambda_{\eta,\alpha,\beta}(x)$. We may substitute this in (5.10) to obtain:

$$\begin{aligned} \Lambda_{\eta,\alpha,\beta}(x) - \langle\zeta,\eta\rangle - \langle\phi,\alpha\rangle - \langle\psi,\beta\rangle + \sigma'\{|\bar{\eta}-\eta|^2 + \|\bar{\alpha}-\alpha\|_2^2 + \|\bar{\beta}-\beta\|_2^2\} \\ \geq \mathrm{i}bid(\bar{\eta},\bar{\alpha},\bar{\beta},\bar{x}) \quad (5.11) \end{aligned}$$

Our notation follows [2] where $\mathrm{i}bid(\bar{\eta},\bar{\alpha},\bar{\beta},\bar{x})$ is used to represent the left hand side of the inequality but with the variables (η,α,β,x) substituted by $(\bar{\eta},\bar{\alpha},\bar{\beta},\bar{x})$. We convert the parameter variables to perturbed arcs using the following change of notation:

$$
\begin{aligned}
x(\cdot) + \alpha(\cdot) &:= u(\cdot) & \dot{x}(\cdot) &:= v_1(\cdot) \\
x(T) + \eta &:= \gamma \\
\bar{x}(\cdot) + \bar{\alpha}(\cdot) &:= \bar{u}(\cdot) & \dot{\bar{x}}(\cdot) &:= \bar{v}_1(\cdot) \\
\bar{x}(T) + \bar{\eta} &:= \bar{\gamma} \\
x(\cdot - \Delta(\cdot)) + \beta(\cdot) &:= w(\cdot) & \dot{x}(\cdot - \Delta(\cdot)) &:= v_2(\cdot) \\
\bar{x}(\cdot - \Delta(\cdot)) + \bar{\beta}(\cdot) &:= \bar{w}(\cdot) & \dot{\bar{x}}(\cdot - \Delta(\cdot)) &:= \bar{v}_2(\cdot)
\end{aligned}
$$

where $v_1(t) := 0$ when $t \in [-\Delta_0, 0]$ and $v_2(t) := 0$ when $t - \Delta(t) < 0$ since we are not concerned with the derivative of the tail $c(\cdot)$, which may not exist. Now, using the new variables, (5.11) becomes

$$
\begin{aligned}
\ell\big(\gamma\big) \; + \; & \int_0^T L\big(t, u(t), w(t), v_1(t), v_2(t)\big)\, dt - \langle \zeta, \gamma - x(T) \rangle \\
& - \int_0^T \{ \langle \phi(t), u(t) - x(t) \rangle + \langle \psi(t), w(t) - x(t - \Delta(t)) \rangle \} \, dt \\
& + \sigma' \{ \|u - x - \bar{u} + \bar{x}\|_2^2 + |\gamma - x(T) - \bar{\gamma} + \bar{x}(T)|^2 \\
& \qquad + \|w - x_\Delta - \bar{w} + \bar{x}_\Delta\|_2^2 \} \geq \mathrm{i}bid(\bar{\gamma}, \bar{u}, \bar{x}, \bar{w}, \bar{v}_1, \bar{v}_2). \qquad (5.12)
\end{aligned}
$$

In an effort to simplify notation, we are using $x_\Delta := x(t - \Delta(t))$ and $\bar{x}_\Delta := \bar{x}(t - \Delta(t))$. Since for any a, b we have $(a + b)^2 \leq 2a^2 + 2b^2$ the square terms in (5.12) can be separated and the result is,

$$
\begin{aligned}
\ell\big(\gamma\big) \; + \; & \int_0^T L\big(t, u(t), w(t), v_1(t), v_2(t)\big)\, dt - \langle \zeta, \gamma - x(T) \rangle \\
& - \int_0^T \{ \langle \phi(t), u(t) - x(t) \rangle + \langle \psi(t), w(t) - x(t - \Delta(t)) \rangle \} \, dt \\
& + 2\sigma' \{ \|u - \bar{u}\|_2^2 + \|x - \bar{x}\|_2^2 + \|x_\Delta - \bar{x}_\Delta\|_2^2 + |\gamma - \bar{\gamma}|^2 \\
& \qquad + |x(T) - \bar{x}(T)|^2 + \|w - \bar{w}\|_2^2 \} \geq \mathrm{i}bid(\bar{\gamma}, \bar{u}, \bar{x}, \bar{w}, \bar{v}_1, \bar{v}_2). \qquad (5.13)
\end{aligned}
$$

In order to eliminate the $x(\cdot)$ terms in the above equation we will use the following estimates:

$$\begin{aligned}
\|x(t)-\bar{x}(t)\|_2 &= \left\|(x(0)-\bar{x}(0))+\int_0^t \{v_1(s)-\bar{v}_1(s)\}\,ds\right\|_2 \\
&\leq \sqrt{T}\,\|v_1(t)-\bar{v}_1(t)\|_1
\end{aligned}$$

$$\begin{aligned}
|x(T)-\bar{x}(T)| &= \left|(x(0)-\bar{x}(0))+\int_0^T \{v_1(s)-\bar{v}_1(s)\}ds\right| \\
&\leq |x(0)-\bar{x}(0)|+\int_0^T |v_1(s)-\bar{v}_1(s)|\,ds \\
&= \|v_1-\bar{v}_1\|_1
\end{aligned}$$

$$\begin{aligned}
\|x_\Delta-\bar{x}_\Delta\|_2 &= \left\|(x(0)-\bar{x}(0))+\int_0^t \{v_2(s)-\bar{v}_2(s)\}[1-\dot{\Delta}(s)]\,ds\right\|_2 \\
&\leq \sqrt{T}(1+K_\Delta)\,\|v_2(t)-\bar{v}_2(t)\|_1\,.
\end{aligned}$$

Substitution of these estimates into (5.13) results in

$$\begin{aligned}
\ell(\gamma)\ +\ &\int_0^T L\big(t,u(t),w(t),v_1(t),v_2(t)\big)\,dt-\langle\zeta,\gamma-x(T)\rangle \\
&-\int_0^T \{\langle\phi(t),u(t)-x(t)\rangle+\langle\psi(t),w(t)-x(t-\Delta(t))\rangle\}\,dt \\
&+\sigma\{\|u-\bar{u}\|_2^2+\|w-\bar{w}\|_2^2+\|v_1-\bar{v}_1\|_1^2+\|v_2-\bar{v}_2\|_1^2+|\gamma-\bar{\gamma}|^2\} \\
&\geq \mathrm{i}bid(\bar{\gamma},\bar{u},\bar{w},\bar{x},\bar{v}_1,\bar{v}_2), \quad (5.14)
\end{aligned}$$

where $\sigma := 2\sigma'(\sqrt{T}+1)(1+K_\Delta)$.

Now, define arcs $p(\cdot)$ and $q(\cdot)$ in $L^2[0,T]$ by

$$p(t)=-\zeta-\int_t^T \phi(s)\,ds \tag{5.15}$$

$$q(t)=-\int_t^T \psi(s)\,ds\,. \tag{5.16}$$

Using integration by parts of we obtain,

$$\begin{aligned}\int_0^T \langle \phi(t), x(t)\rangle\, dt &= \int_0^T \langle \dot{p}(t), x(t)\rangle\, dt \\ &= \langle p(t), x(t)\rangle\Big|_0^T - \int_0^T \langle p(t), v_1(t)\rangle\, dt \\ &= \langle -p(T), x(T)\rangle - \langle p(0), x(0)\rangle - \int_0^T \langle p(t), v_1(t)\rangle\, dt\end{aligned}$$

and

$$\begin{aligned}\int_0^T \langle \psi(t), x(t-\Delta(t))\rangle\, dt &= \int_0^T \langle \dot{q}(t), x(t-\Delta(t))\rangle\, dt \\ &= \langle q(t), x(t-\Delta(t))\rangle\Big|_0^T - \int_0^T \langle q(t), \dot{x}(t-\Delta(t))[1-\dot{\Delta}(t)]\rangle\, dt \\ &= 0 - \langle q(0), x(-\Delta(0))\rangle - \int_0^T \langle q(t)[1-\dot{\Delta}(t)], v_2(t)\rangle\, dt.\end{aligned}$$

Substitution of these into (5.14) and since $x(-\Delta(0)) = c(-\Delta(0))$,

$$\begin{aligned}&\ell\big(\gamma\big) - \langle q(0), c(-\Delta(0))\rangle \;-\; \langle p(0), x(0)\rangle + \langle p(T), \gamma\rangle + \sigma|\gamma - \bar{\gamma}|^2 \\ &\quad + \int_0^T \{L(t, u(t), w(t), v_1(t), v_2(t)) - \langle \dot{p}(t), u(t)\rangle - \langle \dot{q}(t), w(t)\rangle - \langle p(t), v_1(t)\rangle \\ &\qquad\qquad - \langle q(t)[1-\dot{\Delta}(t)], v_2(t)\rangle\}\, dt \\ &\quad + \sigma\{\|u-\bar{u}\|_2^2 + \|w-\bar{w}\|_2^2 + \|v_1-\bar{v}_1\|_1^2 + \|v_2-\bar{v}_2\|_1^2\} \\ &\qquad \geq \mathrm{i}bid(\bar{\gamma}, \bar{u}, \bar{w}, \bar{v}_1, \bar{v}_2, \bar{x})\,. \end{aligned} \tag{5.17}$$

But notice that the terms $\langle q(0), c(-\Delta(0))\rangle$ and $\langle p(0), x(0)\rangle$ are on each side of the inequality, do not depend on $\gamma, , u, w, v_1, v_2$ or x, and that $x(0) = \bar{x}(0)$. So, we may subtract them from each side of the inequality. Finally, we have

$$
\begin{aligned}
&\ell\big(\gamma\big) + \langle p(T), \gamma\rangle + \sigma|\gamma - \bar{\gamma}|^2 \\
&\quad + \int_0^T \{L(t, u(t), w(t), v_1(t), v_2(t)) - \langle \dot{p}(t), u(t)\rangle - \langle \dot{q}(t), w(t)\rangle - \langle p(t), v_1(t)\rangle \\
&\qquad\qquad\qquad\qquad - \langle q(t)[1 - \dot{\Delta}(t)], v_2(t)\rangle\}\, dt \\
&\quad + \sigma\{\|u - \bar{u}\|_2^2 + \|w - \bar{w}\|_2^2 + \|v_1 - \bar{v}_1\|_1^2 + \|v_2 - \bar{v}_2\|_1^2\} \geq \mathrm{i}bid(\bar{\gamma}, \bar{u}, \bar{w}, \bar{v}_1, \bar{v}_2)\,.
\end{aligned}
\tag{5.18}
$$

We know that (5.18) holds as long as $\|\alpha - \bar{\alpha}\|_2$, $|\eta - \bar{\eta}|$, and $\|\beta - \bar{\beta}\|_2$ are each less than δ'. But by definition

$$
\begin{aligned}
\|\alpha - \bar{\alpha}\|_2 &= \|(u - \bar{u}) - (x - \bar{x})\|_2, \\
|\eta - \bar{\eta}| &= |(\gamma - \bar{\gamma}) - (x(T) - \bar{x}(T))|, \qquad \text{and} \\
\|\beta - \bar{\beta}\|_2 &= \|(w - \bar{w}) - (x(t - \Delta(t)) - \bar{x}(t - \Delta(t)))\|_2.
\end{aligned}
$$

Since we bounded $\|x - \bar{x}\|_2$, $\|x_\Delta - \bar{x}_\Delta\|_2$, and $|x(T) - \bar{x}(T)|$ in terms of $\|v_1 - \bar{v}_1\|_1$ and $\|v_2 - \bar{v}_2\|_1$, we only need for $\|u - \bar{u}\|_2$,$\|w - \bar{w}\|_2$,$\|v_1 - \bar{v}_1\|_1$, $\|v_2 - \bar{v}_2\|_1$, and $|\gamma - \bar{\gamma}|$ to be small in order that our inequality (5.18) holds.

We now verify the statements (a) through (d) of the theorem.

(a)We must show that $\mathcal{D}\big(\bar{\gamma}, \bar{u}(\cdot), \bar{w}(\cdot), \bar{v}_1(\cdot), \bar{v}_2(\cdot)\big) < \epsilon$. Notice that

$$
\int_0^T \left|\bar{u}(t) - c(0) - \int_0^t \bar{v}_1(s)\, ds\right|^2 dt = \|\bar{x} - \bar{u}\|_2^2 = \|\bar{\alpha}\|_2^2,
$$

$$
\left|\bar{\gamma} - c(0) - \int_0^T \bar{v}_1(t)\, dt\right| = |\bar{x}(T) - \bar{\gamma}| = |\bar{\eta}|,
$$

$$
\int_0^T |v_2(t) - v_1(t - \Delta(t))|\; dt = \int_0^T |\dot{x}(t - \Delta(t)) - \dot{x}(t - \Delta(t))|\; dt = 0,
$$

and

$$\int_0^T \left| \bar{w}(t) - \xi(t)\left[c(t-\Delta(t)) - x(0)\right] - x(0) - \int_0^t \bar{v}_2(s)[1-\dot{\Delta}(s)]\,ds \right|^2 dt$$
$$= \int_0^T \left| \bar{w}(t) - x(0) - \int_0^t \bar{v}_2(s)[1-\dot{\Delta}(s)]\,ds \right|^2 dt$$
$$= \|\bar{w}(t) - \bar{x}(t-\Delta(t))\|_2 = \|\beta\|_2^2.$$

But, $\|\bar{\alpha}\|_2 + |\bar{\eta}| + \|\bar{\beta}\|_2 < \epsilon$ by our original choice of $\bar{\alpha}$, $\bar{\eta}$, and $\bar{\beta}$, so that statement (a) is satisfied. Now, lets prove (b).

(b)To show that the minimum value of $\Gamma\big(\bar{\gamma}, \bar{u}(\cdot), \bar{w}(\cdot), \bar{v}_1(\cdot), \bar{v}_2(\cdot)\big)$ is within ϵ of the minimum value in (4.3) we observe that min $\Gamma\big(\bar{\gamma}, \bar{u}(\cdot), \bar{w}(\cdot), \bar{v}_1(\cdot), \bar{v}_2(\cdot)\big) = V(\bar{\eta}, \bar{\alpha}, \bar{\beta})$ and that $V(0,0,0)$ is the minimum for equation (4.3). Therefore, since $|V(\bar{\eta}, \bar{\alpha}, \bar{\beta}) - V(0,0,0)| < \epsilon$, it follows that (b) holds. Statement (d) is satisfied by setting $u = \bar{u}, v_1 = \bar{v}_1, v_2 = \bar{v}_2, w = \bar{w}$ in (5.18). Thus, we have only left to show that (c) holds.

Part (c) states that for almost all $t \in [0,T]$, the map

$$\begin{aligned}(u, w, v_1, v_2) \mapsto L(t, u, w, v_1, v_2) &- \langle \dot{p}(t), u\rangle - \langle \dot{q}(t), w\rangle - \langle p(t), v_1\rangle \\ &- \langle q(t)[1-\dot{\Delta}(t)], v_2\rangle \\ &+ \sigma\Big\{|u-\bar{u}(t)|^2 + |w-\bar{w}(t)|^2 + |v_1-\bar{v}_1(t)|^2 + |v_2-\bar{v}_2(t)|^2\Big\}\end{aligned}$$

is minimized at $(u, w, v_1, v_2) = \big(\bar{u}(t), \bar{w}(t), \bar{v}_1(t), \bar{v}_2(t)\big)$. To prove this let f_t denote the function

$$\begin{aligned}f_t(u, w, v_1, v_2) = L(t, u, w, v_1, v_2) - \langle \dot{p}(t), u\rangle - \langle \dot{q}(t), w\rangle - \langle p(t), v_1\rangle - \langle q(t)[1-\dot{\Delta}(t)], v_2\rangle \\ + \sigma\{|u-\bar{u}(t)|^2 + |w-\bar{w}(t)|^2 + |v_1-\bar{v}_1(t)|^2 + |v_2-\bar{v}_2(t)|^2\}. \quad (5.19)\end{aligned}$$

Hypotheses (H2) and (H3) imply that f_t attains a minimum for almost every t. We will prove part (c) by establishing that

$$f_t(\bar{u}(t), \bar{w}(t), \bar{v}_1(t), \bar{v}_2(t)) = \min_{(u,w,v_1,v_2)\in R^{4n}} f_t(u, w, v_1, v_2).$$

In fact, it is sufficient to show that for each $r > 0$ the set

$$A(r) := \left\{ t \in [0,T] : f_t(\bar{u}(t), \bar{w}(t), \bar{v}_1(t), \bar{v}_2(t)) \geq \min_{(u,w,v_1,v_2)\in R^{4n}} f_t(u,w,v_1,v_2) + r \right\}$$

has measure zero.

Suppose not, then there exists an $r > 0$ with $m(A(r)) > 0$. Let B_i be a decreasing sequence of subsets of $A(r)$ such that for each i one has

$$\frac{1}{i} m(A(r)) < m(B_i) < \frac{3}{i} m(A(r)).$$

Choose $(u'(\cdot), w'(\cdot), v_1'(\cdot), v_2'(\cdot))$ measurable such that for almost all t, $(u'(t), w'(t), v_1'(t), v_2'(t))$ minimizes f_t. The existence of this function follows from standard measurable selection results ([25], pp. 157-207). To complete our proof we must first show that u',w', $v_1' - \bar{v}_1$, and $v_2' - \bar{v}_2$ are in $L^2[0,T]$. Consider the inequality

$$f_t(u'(t), w'(t), v_1'(t), v_2'(t)) \leq f_t(\bar{u}(t), \bar{w}(t), \bar{v}_1(t), \bar{v}_2(t)),$$

which by our choice of $(u'(\cdot), w'(\cdot), v_1'(\cdot), v_2'(\cdot))$ holds for almost all t. Let b be the lower bound for the function L. After simplifying, we obtain from our definition of f_t in (5.19),

$$\begin{aligned}
L(t, u'(t), w'(t), v_1'(t), v_2'(t)) - \langle \dot{p}(t), u'(t)\rangle - \langle \dot{q}(t), w'(t)\rangle \\
+ \langle p(t), v_1'(t)\rangle + \langle q(t)[1-\dot{\Delta}(t)], v_2'(t)\rangle \\
+ \sigma\{|u'(t)-\bar{u}(t)|^2 + |w'(t)-\bar{w}(t)|^2 + |v_1'(t)-\bar{v}_1(t)|^2 + |v_2'(t)-\bar{v}_2(t)|^2\} \\
\leq L(t, \bar{u}(t), \bar{w}(t), \bar{v}_1(t), \bar{v}_2(t)) - \langle \dot{p}(t), \bar{u}(t)\rangle \\
- \langle \dot{q}(t), \bar{w}(t)\rangle + \langle p(t), v_1'(t)\rangle + \langle q(t)[1-\dot{\Delta}(t)], v_2'(t)\rangle . \quad (5.20)
\end{aligned}$$

Replacing $L(t, u'(t), w'(t), v_1'(t), v_2'(t))$ by its lowerbound b and rearranging we obtain,

$$\begin{aligned}
&\sigma\{|u'(t)-\bar{u}(t)|^2+|w'(t)-\bar{w}(t)|^2+|v_1'(t)-\bar{v}_1(t)|^2+|v_2'(t)-\bar{v}_2(t)|^2\} \\
&\qquad \le \langle \dot{p}(t), u'(t)-\bar{u}(t)\rangle + \langle \dot{q}(t), w'(t)-\bar{w}(t)\rangle + \langle p(t), v_1'(t)-\bar{v}_1(t)\rangle \\
&\qquad\qquad + \langle q(t), v_2'(t)-\bar{v}_2(t)\rangle + L(t,\bar{u}(t),\bar{w}(t),\bar{v}_1(t),\bar{v}_2(t)) - b\,. \quad (5.21)
\end{aligned}$$

Set

$$\begin{aligned}
W(t) &= (u'(t)-\bar{u}(t), w'(t)-\bar{w}(t), v_1'(t)-\bar{v}_1(t), v_2'(t)-\bar{v}_2(t))\ , \\
g(t) &= \Big(\dot{p}(t), \dot{q}(t), p(t), q(t)[1-\dot{\Delta}(t)]\Big)\ , \\
k(t) &= L(t,\bar{u}(t),\bar{w}(t),\bar{v}_1(t),\bar{v}_2(t)) - b\,.
\end{aligned}$$

Then inequality (5.21) becomes $\sigma\,|W|^2 - \langle g(t), W(t)\rangle - k(t) \le 0$ and since for any inner product $|\langle \alpha, \beta\rangle| \le \|\alpha\|\|\beta\|$, by the quadratic formula

$$|W|^2 \le \frac{|g| + \sqrt{|g|^2 + 4k(t)\sigma}}{2\sigma}.$$

Thus, $(2\sigma\,|W|^2 - |g|)^2 \le |g|^2 + 4k(t)\sigma = |g|^2 + 4|k(t)|\sigma$, because by definition $k(t) \ge 0$ for almost all t. Furthermore, since $g(\cdot) \in L^2$ and $k(\cdot) \in L^1$, it follows that the integral of $(2\sigma\,|W|^2 - |g|)^2$ is bounded and thus, $W \in L^2$, proving that u', w',$v_1' - \bar{v}_1$, and $v_2' - \bar{v}_2$ are in $L^2[0,T]$. Now, since for each $j = 1,2$; $v_j' - \bar{v}_j \in L^2[0,T]$ and $\bar{v}_j \in L^1[0,T]$, it follows that each v_1' and v_2' is also in $L^1[0,T]$.

We continue now with the proof of (c). Define a family of functions $(u_i, w_i, v_{1i}, v_{2i}) \in L^2[0,T] \times L^2[0,T] \times L^1[0,T] \times L^1[0,T]$ as:

$$(u_i(t), w_i(t), v_{1i}(t), v_{2i}(t)) = \begin{cases} (u'(t), w'(t), v_1'(t), v_2'(t)), & \text{if } t \in B_i; \\ (\bar{u}(t), \bar{w}(t), \bar{v}_1(t), \bar{v}_2(t)), & \text{else}\,. \end{cases} \quad (5.22)$$

Then,

$$\lim_{i\to\infty} \|(u_i, w_i, v_{1i}, v_{2i}) - (\bar{u}, \bar{w}, \bar{v}_1, \bar{v}_2)\|_2 = 0,$$

since as $i \to \infty$, $m(B_i) \to 0$. From the definition of (u', w', v_1', v_2') and since $m(A(r)) > 0$, it follows that

$$\int_0^T f_t(u_i, w_i, v_{1i}, v_{2i})\, dt < \int_0^T f_t(\bar{u}, \bar{w}, \bar{v}_1, \bar{v}_2)\, dt.$$

But for i sufficiently large this contradicts equation (5.18), setting $\gamma = \bar{\gamma}, u = u_i, w = w_i$, $v_1 = v_{1i}$, and $v_2 = v_{2i}$. Hence, it must be that $m(A(r)) = 0$, which proves (c) and completes our proof. $\square$

An immediate consequence of statement (c) in Theorem 5.1.1 is that for almost all $t \in [0, T]$, the inclusion

$$\big(\dot{p}(t), \dot{q}(t), p(t), \bar{q}(t)\big) \in \partial_P L\big(t, \cdot, \cdot, \cdot, \cdot\big)\big(\bar{u}(t), \bar{w}(t), \bar{v}_1(t), \bar{v}_2(t)\big) \tag{5.23}$$

holds, which is a rudimentary form of the Euler-Lagrange equation. Here, $\bar{q}(t) = q(t)[1 - \dot{\Delta}(t)]$ and $\partial_P L(t, \cdot, \cdot, \cdot, \cdot)\big(\bar{u}(t), \bar{w}(t), \bar{v}_1(t), \bar{v}_2(t)\big)$ refers to the proximal subgradient of L with respect to the (x, y, v_1, v_2) variables evaluated at $\big(\bar{u}(t), \bar{w}(t), \bar{v}_1(t), \bar{v}_2(t)\big)$. An immediate consequence of (d) is that

$$-p(T) \in \partial_P \ell(\bar{\gamma}), \tag{5.24}$$

which is a forerunner of the transversality condition.

The inclusion (5.23) can be dualized and caste in Hamiltonian form as well. Define $H_\sigma : [0, T] \times I\!R^n \times I\!R^n \times I\!R^n \times I\!R^n \to I\!R$ by

$$H_\sigma(t, x, y, p, q) := \sup_{(v,w)\in R^{2n}} \Big\{ \langle p, v\rangle + \langle q, w\rangle - L(t, x, y, v, w) - \sigma\big\{|v - \bar{v}_1(t)|^2 + |w - \bar{v}_2(t)|^2\big\}\Big\}.$$

The case $\sigma = 0$ is the Hamiltonian defined in (4.8) and is denoted simply as H. Further consequences of (5.23) include the following Proposition holding almost everywhere on $[0, T]$:

Proposition 5.1.3. *The following follow from (c):*

(i) $\big(\dot p(t), \dot q(t)\big) \in \partial_P(-H_\sigma)\big(t, \cdot, \cdot, p(t), q(t)[1-\dot\Delta(t)]\big)\big(\bar u(t), \bar w(t)\big)$

(ii) $\big(\bar v_1(t), \bar v_2(t)\big) \in \partial_P H_0\big(t, \bar u(t), \bar w(t), \cdot, \cdot\big)\big(p(t), q(t)[1-\dot\Delta(t)]\big)$

If H_σ is locally Lipschitz in (x, y, p, q) then we also have,

(iii) $\big(-\dot p(t), -\dot q(t), \bar v_1(t), \bar v_2(t)\big) \in \partial_P H_\sigma\big(t, \cdot, \cdot, \cdot, \cdot\big)(\bar u(t), \bar w(t), p(t), q(t)[1-\dot\Delta(t)]),$

where in the above equations the proximal subgradient is taken with respect to the $(\cdot)$ variable.

Proof. Choose $t \in [0, T]$ so that (c) holds. We first prove (ii). From (c) we have that

$$\begin{aligned} L(t, u, w, v_1, v_2) - \langle \dot p(t), u\rangle - \langle \dot q(t), w\rangle - \langle p(t), v_1\rangle - \langle q(t)[1-\dot\Delta(t)], v_2\rangle \\ + \sigma\Big\{|u-\bar u(t)|^2 + |w-\bar w(t)|^2 + |v_1 - \bar v_1(t)|^2 + |v_2 - \bar v_2(t)|^2\Big\} \\ \geq ibid(\bar u(t), \bar w(t), \bar v_1(t), \bar v_2(t)). \end{aligned} \tag{5.25}$$

Setting $u = \bar u(t)$ and $w = \bar w(t)$, in (5.25) we obtain

$$\begin{aligned} L(t, \bar u(t), \bar w(t), v_1, v_2) - \langle p(t), v_1\rangle - \langle q(t)[1-\dot\Delta(t)], v_2\rangle + \sigma\big\{|v_1 - \bar v_1(t)|^2 + |v_2 - \bar v_2(t)|^2\big\} \\ \geq L(t, \bar u(t), \bar w(t), \bar v_1(t), \bar v_2(t)) - \langle p(t), \bar v_1(t)\rangle - \langle q(t)[1-\dot\Delta(t)], \bar v_2(t)\rangle, \end{aligned}$$

setting $v_2 = \bar v_2(t)$ we have that $v_1 = \bar v_1(t)$ maximizes

$$\langle p(t), v_1\rangle - L(t, \bar u(t), \bar w(t), v_1, \bar v_2(t)) - \sigma\left|v_1 - \bar v_1(t)\right|^2, \tag{5.26}$$

and setting $v_1 = \bar v_1(t)$ we have that $v_2 = \bar v_2(t)$ maximizes

$$\langle q(t)[1-\dot\Delta(t)], v_2\rangle - L(t, \bar u(t), \bar w(t), \bar v_1(t), v_2) - \sigma\left|v_2 - \bar v_2(t)\right|^2. \tag{5.27}$$

But, the above, (5.26)and (5.27), imply that

$$0 \in \partial_P \left\{ -\langle p(t), \cdot \rangle + L(t, \bar{u}(t), \bar{w}(t), \cdot, \bar{v}_2(t)) + \sigma \left| \cdot - \bar{v}_1(t) \right|^2 \right\} (\bar{v}_1(t))$$

and

$$0 \in \partial_P \left\{ -\langle q(t)[1 - \dot{\Delta}(t)], \cdot \rangle + L(t, \bar{u}(t), \bar{w}(t), \bar{v}_1(t), \cdot) + \sigma \left| \cdot - \bar{v}_2(t) \right|^2 \right\} (\bar{v}_2(t)) .$$

But, since $\langle p(t), v_1 \rangle$ and $\sigma \left| v_1 - \bar{v}_1(t) \right|^2$ are twice continuously differentiable in v_1, as are $\langle q(t), v_2 \rangle$ and $\sigma \left| v_2 - \bar{v}_2(t) \right|^2$ in v_2, by a well known result in non-smooth analysis [3] we obtain

$$p(t) + 2\sigma \left| v_1 - \bar{v}_1(t) \right| \in \partial_P L(t, \bar{u}(t), \bar{w}(t), \cdot, \bar{v}_2(t))(\bar{v}_1(t))$$

and

$$q(t)[1 - \dot{\Delta}(t)] + 2\sigma \left| v_2 - \bar{v}_2(t) \right| \in \partial_P L(t, \bar{u}(t), \bar{w}(t), \bar{v}_1(t), \cdot)(\bar{v}_2(t)).$$

After evaluating the left hand side of the previous two inclusions at $v_1 = \bar{v}_1(t)$,and $v_2 = \bar{v}_2(t)$ respectively, we have

$$\big(p(t), q(t)[1 - \dot{\Delta}(t)]\big) \in \partial_P L(t, \bar{u}(t), \bar{w}(t), \cdot, \cdot)(\bar{v}_1(t), \bar{v}_2(t)). \tag{5.28}$$

By convex analysis we know that $L(t, \bar{u}(t), \bar{w}(t), v_1, v_2)$ is the conjugate of $H_0(t, \bar{u}(t), \bar{w}(t), p, q)$ so that (5.28) holds if and only if $(\bar{v}_1(t), \bar{v}_2(t)) \in \partial_P H_0(t, \bar{u}(t), \bar{w}(t), p(t), q(t)[1 - \dot{\Delta}(t)])$. Thus, (ii) holds. Now, to show that (i) holds, consider again the original inequality (5.25) implied by (c). Rearranging we obtain

$$\begin{aligned}
\sup_{(v_1,v_2)\in R^{2n}} & \{ \langle p(t), v_1 \rangle + \langle q(t)[1 - \dot{\Delta}(t)], v_2 \rangle - L(t, u, w, v_1, v_2) \\
& \qquad - \sigma \big\{ \left| v_1 - \bar{v}_1(t) \right|^2 + \left| v_2 - \bar{v}_2(t) \right|^2 \big\} \} \\
& \leq \langle \dot{p}(t), \bar{u}(t) - u \rangle + \langle \dot{q}(t), \bar{w}(t) - w \rangle + \sigma \left\{ \left| u - \bar{u}(t) \right|^2 + \left| w - \bar{w}(t) \right|^2 \right\} \\
& + \langle p(t), \bar{v}_1(t) \rangle + \langle q(t)[1 - \dot{\Delta}(t)], \bar{v}_2(t) \rangle - L(t, \bar{u}(t), \bar{w}(t), \bar{v}_1(t), \bar{v}_2(t)).
\end{aligned} \tag{5.29}$$

Notice that we may take the supremum of the left hand side without changing the inequality since the right hand side of the inequality is independent of v_1 and v_2. Now, (5.29) is equivalent to

$$H_\sigma\big(t, u, w, p(t), q(t)[1-\dot{\Delta}(t)]\big) \leq H_\sigma\big(t, \bar{u}(t), \bar{w}(t), p(t), q(t)[1-\dot{\Delta}(t)]\big) + \sigma\left\{|u-\bar{u}(t)|^2 + |w-\bar{w}(t)|^2\right\} + \langle \dot{p}(t), \bar{u}(t) - u\rangle + \langle \dot{q}(t), \bar{w}(t) - w\rangle,$$

which, finally, is equivalent to

$$(-H_\sigma)(t, u, w, p(t), q(t)[1-\dot{\Delta}(t)]) - (-H_\sigma)(t, \bar{u}(t), \bar{w}(t), p(t), q(t)[1-\dot{\Delta}(t)]) + \sigma\left\{|u-\bar{u}(t)|^2 + |w-\bar{w}(t)|^2\right\} \geq \langle \dot{p}(t), u - \bar{u}(t)\rangle + \langle \dot{q}(t), w - \bar{w}(t)\rangle. \quad (5.30)$$

By the proximal subgradient inequality [3] (5.30) implies (i). In order to prove (iii) suppose that H_σ is locally Lipschitz in (x, y, p, q), then for any $\mu := (\mu_1, \mu_2, \mu_3, \mu_4)$ in $I\!R^{4n}$ the directional derivative $H^0_\sigma(t, \bar{u}(t), \bar{w}(t), p(t), q(t)[1-\dot{\Delta}(t)]; \mu)$ majorizes

$$\limsup_{\lambda\to 0} \frac{1}{\lambda} \cdot \big\{\mathrm{H}_\sigma(t, \bar{u}(t), \bar{w}(t), p(t) + \lambda\mu_3, q(t)[1-\dot{\Delta}(t)] + \lambda\mu_4) - H_\sigma(t, \bar{u}(t) - \lambda\mu_1, \bar{w}(t) - \lambda\mu_2, p(t), q(t)[1-\dot{\Delta}(t)])\big\}$$

$$\geq \limsup_{\lambda\to 0} \frac{1}{\lambda} \cdot \{\mathrm{H}_\sigma(t, \bar{u}(t), \bar{w}(t), p(t), q(t)[1-\dot{\Delta}(t)]) + \lambda\langle \mu_3, \bar{v}_1(t)\rangle + \lambda\langle \mu_4, \bar{v}_2(t)\rangle - H_\sigma(t, \bar{u}(t) - \lambda\mu_1, \bar{w}(t) - \lambda\mu_2, p(t), q(t)[1-\dot{\Delta}(t)])\}$$

$$\geq \limsup_{\lambda\to 0} \frac{1}{\lambda} \cdot \big\{\langle \dot{p}(t), -\lambda\mu_1\rangle + \langle \dot{q}(t), -\lambda\mu_2\rangle - \sigma\big\{|\lambda\mu_1|^2 + |\lambda\mu_2|^2\big\} + \lambda\langle \mu_3, \bar{v}_1(t)\rangle + \lambda\langle \mu_4, \bar{v}_2(t)\rangle\big\} \quad \text{by (5.30)}$$

$$= -\langle \dot{p}(t), \mu_1\rangle - \langle \dot{q}(t), \mu_2\rangle + \langle \mu_3, \bar{v}_1(t)\rangle + \langle \mu_4, \bar{v}_2(t)\rangle.$$

Thus, $(-\dot{p}(t), -\dot{q}(t), \bar{v}_1(t), \bar{v}_2(t)) \in \partial_P H_\sigma(t, \cdot, \cdot, \cdot, \cdot)(\bar{u}(t), \bar{w}(t), p(t), q(t)[1-\dot{\Delta}(t)])$ so that (iii) is proved. $\square$

5.2 Necessary Conditions for the Neutral Problem of Bolza

The decoupling technique is a powerful tool for generating necessary conditions, as may be surmised by observing (5.23), (5.24), and the results from the previous Proposition. See [31] for state-of-the-art results for systems without time delays. We now add the following additional hypothesis:

(H4) $\ell(\gamma)$ is locally Lipschitz for each $\gamma \in B(x(T), \epsilon)$ for some $\epsilon > 0$ as in Theorem 1 and L is locally Lipschitz on variables (x, y, v_1, v_2). Also, for all $(\Omega, \Upsilon, \Theta, \Phi) \in \partial_P L(t, x, y, v, w)$, we assume the following growth conditions:

$$\begin{aligned} \|\Omega\| &\leq k_1(t)\|\Theta\| + k_2(t) \\ \|\Upsilon\| &\leq k_3(t)\|\Phi\| + k_4(t) \end{aligned}$$

where $k_1(\cdot), k_2(\cdot), k_3(\cdot)$,and $k_4(\cdot)$ are integrable and positive for all $t \in [0, T]$.

Assuming (H4), we are now able to conclude that there exist arcs $p(\cdot)$ and $q(\cdot)$ simultaneously satisfying both the Eulerian and Hamiltonian relations. This is the same result obtained by Clarke in [2]. We prove this result in the following theorem.

Theorem 5.2.1. *Assume (H1)-(H4). Let $x(\cdot)$ solve the Bolza problem (4.4), then there exist arcs $p(\cdot)$ and $q(\cdot)$ with ($\bar{q}(t) = q(t)[1 - \dot{\Delta}(t)]$)satisfying the following a.e.:*

(1) $(\dot{p}(t), \dot{q}(t), p(t), \bar{q}(t)) \in \partial_C L(t, \cdot, \cdot, \cdot, \cdot)(x(t), x(t - \Delta(t)), \dot{x}(t), \dot{x}(t - \Delta(t)))$

(2) $(-\dot{p}(t), -\dot{q}(t), \dot{x}(t), \dot{x}(t - \Delta(t))) \in \partial_C H(t, \cdot, \cdot, \cdot, \cdot)(x(t), x(t - \Delta(t)), p(t), \bar{q}(t))$

(3) $-p(T) \in \partial_L \ell(x(T))$

Proof. Assume the hypotheses of the theorem. Modify L by $\widetilde{L}(t, x, y, v_1, v_2) = L(t, x, y, v_1, v_2) + |v_1 - \dot{x}(t)|^2 + |v_2 - \dot{x}(t - \Delta(t))|^2$. Applying Theorem 1 for a sequence $\epsilon_i \searrow 0$ as $i \to \infty$ and since $\widetilde{H}_\sigma$ for $\widetilde{L}$ is Lipschitz in (x, y, p, q) we have

(4.1) $(\dot{p}_i(t), \dot{q}_i(t), p_i(t), q_i(t)[1-\dot{\Delta}(t)]) \in \partial_P \widetilde{L}(t, \bar{u}_i(t), \bar{w}_i(t), \bar{v}_{1i}(t), \bar{v}_{2i}(t))$ a.e.

(4.2) $-p_i(T) \in \partial_P \ell(\bar{\gamma}_i)$

(4.3) $(-\dot{p}_i(t), -\dot{q}_i(t), \bar{v}_{1i}(t), \bar{v}_{2i}(t)) \in \partial_P \widetilde{H}_\sigma(t, \bar{u}_i(t), \bar{w}_i(t), p_i(t), q_i(t)[1-\dot{\Delta}(t)])$ a.e.,

where (4.1) and (4.2) follow by setting (c) and (d), from Theorem 1, respectively to zero at the minimizers $(t, \bar{u}(t), \bar{w}(t), \bar{v}_1(t), \bar{v}_2(t))$, $\bar{\gamma}$, and (4.3) follows from the previous Proposition. Using a result from Loewen and Rockafellar [13], we conclude that $\partial_P \widetilde{H}_\sigma(t, \bar{u}_i(t), \bar{w}_i(t), p_i(t), q_i(t)[1-\dot{\Delta}(t)])$ is contained for almost every t in $\partial_P \widetilde{H}(t, \bar{u}_i(t), \bar{w}_i(t), p_i(t), q_i(t)[1-\dot{\Delta}(t)])$. Thus, (4.3) holds with $\sigma = 0$. Conclusion (b) of Theorem 1, using $\widetilde{L}$, says that

$$\Big| \ell(\bar{\gamma}) + \int_0^T \{L(t, \bar{u}_i(t), \bar{w}_i(t), \bar{v}_i(t)) + |\bar{V}_i(t)) - \dot{X}(t)|^2\}\, dt \\ - \ell(x(T)) - \int_0^T L(t, x(t), x(t-\Delta(t)), \dot{x}(t), \dot{x}(t-\Delta(t)))dt \Big| < \epsilon_i. \quad (5.31)$$

Here, we let $V_i(t) := (v_{1i}(t), v_{2i}(t))$ and $\dot{X}(t) := (\dot{x}(t), \dot{x}(t-\Delta(t)))$ for simplicity. We may also write the above inequality (5.31) with L equal to our original L obtaining,

$$\Big| \ell(\bar{\gamma}) + \int_0^T L(t, \bar{u}_i(t), \bar{w}_i(t), \bar{v}_{1i}(t), \bar{v}_{2i}(t))\, dt \\ - \ell(x(T)) - \int_0^T L(t, x(t), x(t-\Delta(t)), \dot{x}(t), \dot{x}(t-\Delta(t)))dt \Big| < \epsilon_i. \quad (5.32)$$

Combining (5.31) and (5.32) we can observe that $\int_0^T |\bar{V}_i(t)) - \dot{X}(t)|^2\, dt \mapsto 0$ so that each $v_{1i} \mapsto \dot{x}(t)$ and $v_{2i}(t) \mapsto \dot{x}(t-\Delta(t))$ strongly in $L^2[0,T]$. From part (a) of Theorem 1, we have that $\mathcal{D}(\bar{\gamma}, \bar{u}_i(\cdot), \bar{w}_i(\cdot), \bar{v}_{1i}(\cdot), \bar{v}_{2i}(\cdot)) < \epsilon_i$ so that in particular

$$\int_0^T \left| \bar{u}_i(t) - x(0) - \int_0^t \bar{v}_{1i}(s)\, ds \right|^2 dt < \epsilon_i. \quad (5.33)$$

But, since $\bar{v}_{1i}(t) \to \dot{x}(t)$ in $L^2[0,T]$, (5.33) implies that $\bar{u}_i(t) \to x(t)$. Moreover, we can utilize the definition of $\mathcal{D}$ in a similar way to see that $\bar{w}_i(t) \to x(t-\Delta(t))$

for $t \in [0,T]$ and that $\gamma_i \mapsto x(T)$. Choose i large enough so that $\bar{u}_i(t) \to x(t)$, $\bar{w}_i(t) \to x(t-\Delta(t))$, $\bar{V}_i(t) \to \dot{X}(t)$, and $\gamma_i \to x(T)$ almost everywhere. Since $\ell(\cdot)$ is locally Lipschitz $\ell(\gamma_i) \to \ell(x(T))$. Combining this together with (5.32) we obtain that

$$\int_0^T L(t, \bar{u}_i(t), \bar{w}_i(t), \bar{v}_{1i}(t), \bar{v}_{2i}(t))\, dt \to \\ \int_0^T L(t, x(t), x(t-\Delta(t)), \dot{x}(t), \dot{x}(t-\Delta(t)))\, dt.$$

The locally Lipschitz condition of $\ell(\cdot)$ in (H4) implies that $\{p_i(T)\}_{i=1}^{\infty} < K_\ell$ for some $0 < K_\ell < \infty$ so that there exists a subsequence (without relabeling) $\{p_i(T)\} \to p(T)$, for some absolutely continuous arc $p(\cdot)$. Hence, by a weak sequential compactness theorem ([3],p. 150), we can pass to the limit in (4.2) to obtain that $-p(T) \in \partial_L \ell(x(T))$, proving (3). We have left to show that (1) and (2) hold.

Given (H4) and (4.1) we have that,

$$\begin{aligned} \|\dot{p}_i(t)\| &\le k_1(t)\big\|p_i(t) + 2|v_{1i}(t) - \dot{x}(t)|\big\| + k_2(t) \\ &\le k_1(t)\|p_i(t)\| + 2k_1(t)\|v_{1i}(t) - \dot{x}(t)\| + k_2(t). \end{aligned} \tag{5.34}$$

We use Grownwall's inequality to proceed with our estimate. For $t \in [0,T]$

$$\begin{aligned} \|p_i(t)\| &\le \|p_i(T)\| \mathrm{e}^{k_1(t)(t-T)} + \int_t^T \mathrm{e}^{k_1(t)(t-s)} \big(k_2(s) + 2k_1(s)\|v_{1i}(s) - \dot{x}(s)\|\big)\, ds \\ &\le \|p_i(T)\| + T \cdot \big\{\|k_2\| + 2\|k_1\|\|v_{1i} - \dot{x}\|\big\}. \end{aligned}$$

Since $\|p_i(T)\| = \zeta_i$ is bounded by $p(T)$, the functions k_1, k_2 are in $L^1[0,T]$, and by choosing i large $\|v_{1i} - \dot{x}\|$ can be made arbitrarily small, we have found a bound, call it $\mathcal{M}$, for $\|p_i(t)\|$. We can use a similar argument to show that given (H4)

$$\|q_i(t)\| \le T\{\|k_4\| + 2(K_\Delta + 1)\|k_3\|\|v_{2i} - \dot{x}(t-\Delta(t))\|\}.$$

Hence, the sequences $\{p_i(\cdot)\}$, $\{q_i(\cdot)\}$, $\{\dot{p}_i(\cdot)\}$, and $\{\dot{q}_i(\cdot)\}$ each have a subsequence which converges weakly to some arcs $p(\cdot), q(\cdot), \dot{p}(\cdot), \dot{q}(\cdot)$ respectively. In the case of the p_i's, we use a subsequence of our original subsequence of arcs p_i converging to $p(T)$. We then use a weak sequential compactness theorem to pass to the limit on (4.1) and (4.3) obtaining

$$(4.1') \quad (\dot{p}(t), \dot{q}(t), p(t), \bar{q}(t)) \in \partial_C \widetilde{L}(t, \bar{x}(t), \bar{x}(t-\Delta(t)), \dot{x}(t), \dot{x}(t-\Delta(t))) \text{ a.e.}$$

$$(4.3') \quad (-\dot{p}(t), -\dot{q}(t), \dot{x}(t), \dot{x}(t-\Delta(t))) \in \partial_C \widetilde{H}(t, \bar{x}(t), \bar{x}(t-\Delta(t)), p(t), \bar{q}(t)) \text{ a.e.}$$

with $\bar{q}(t) = q(t)[1 - \dot{\Delta}(t)]$. Using, again, a result from Loewen and Rockafellar [13] in order to remove the tilde from (4.1′) and (4.3′) we have completed the proof of the theorem. □

We have seen thus: Given $\bar{x}(\cdot)$ an optimal solution to (4.4), the additional term $|V - \dot{\bar{X}}(t)|^2$ is added to L in order to single out $\bar{x}(\cdot)$ as the unique optimal solution, which helps validate ensuing limiting arguments without affecting the necessary conditions themselves. The decoupling principle is then applied sequentially with $\epsilon = \epsilon_i$, where $\epsilon_i \to 0$. Necessary conditions will follow provided passage to the limit of all the functions produced in the Theorem can be justified. This requires extra hypotheses beyond (H1)-(H3), but nonetheless the role of these additional hypotheses can be interpreted directly as being the needed requirements for the limiting arguments to be carried out. Furthermore, both Euler-Lagrange and Hamiltonian formulations were produced by the same adjoint arc, since the limiting (5.23) and (i)-(iii) in Proposition 5.1.3 can be done simultaneously.

5.3 Decoupling and the State Time Delay Case

We may use the decoupling principle, analogous to the previous section, to obtain necessary conditions for the generalized problem of Bolza with time delay in the state variable only [21]. We refer to this problem as the Delayed Problem of Bolza or (DPB). The proofs in this section resemble those in the neutral case. We now modify our integral functional as follows

$$\Lambda\big(x(\cdot)\big) := \ell\big(x(T)\big) \ + \int_0^T L\big(t, x(t), x(t-\Delta(t)), \dot{x}(t)\big)\, dt. \tag{5.35}$$

The generalized Bolza problem with time delay is the following optimization problem:

$$\text{minimize } \Lambda\big(x(\cdot)\big) \tag{5.36}$$

over $x(\cdot) \in \mathcal{AC}[0,T]$, and where $x(t) = c(t)$ for almost all $t \in [-\Delta_0, 0]$. Again, the delay is a function of bounded variation $\Delta : [0,T] \to [0, \Delta_0]$ with $|\dot{\Delta}(t)| \leq K_\Delta$, $\Delta(0) = 0$, and where $\Delta_0 > 0$ is a fixed constant. And, $c : [-\Delta_0, 0] \to I\!R^n$ is in $L^2[-\Delta_0, 0]$.

We modify our assumptions to fit this case.

(H1) ℓ is lower semicontinuous and bounded below;

(H2) $L(t,x,y,v)$ is lower semicontinuous in (x,y,v), is $\mathcal{L}\times\mathcal{B}$-measurable on $[0,T]\times I\!R^{3n}$, and is convex in v.

(H3) There exists a function $\theta : [0,\infty) \to I\!R$ satisfying $\lim_{r\to\infty} \theta(r)/r = \infty$ and so that

$$L\big(t,x,y,v\big) \geq \theta\big(|v|\big) \qquad \text{for all } v \in I\!R^n.$$

We treat the presence of the delayed arc $w(t) := x(t - \Delta(t))$ as an additional constraint, related by

$$w(t) = \xi(t)\,[c(t-\Delta(t)) - x_0] + x_0 + \int_0^{\rho(t)} v(s)\, ds \tag{5.37}$$

for $t \in [0,T]$, where

$$\xi(t) = \begin{cases} 1, & \text{if } t - \Delta(t) < 0; \\ 0 & \text{otherwise} \end{cases} \tag{5.38}$$

and $\rho(t) = \max\{0, t-\Delta(t)\}$. Note that (5.37) simply defines $w(t)$ as equal to the initial tail when $t - \Delta(t) \in [-\Delta_0, 0]$ and equal to $x(t - \Delta(t))$ for $t - \Delta(t)$ in $[0,T]$.

Let $\mathcal{X} := I\!R^n \times L^2[0,T] \times L^2[0,T] \times L^1[0,T]$. A new Bolza-type functional $\Gamma : \mathcal{X} \to (-\infty, \infty]$, which is similar to (5.35), is defined by

$$\Gamma\Big(\gamma, u(\cdot), w(\cdot), v(\cdot)\Big) := \ell\big(\gamma\big) \; + \; \int_0^T L\big(t, u(t), w(t), v(t)\big)\, dt. \tag{5.39}$$

It is not difficult to show that problem (5.36) is equivalent to minimizing Γ over $\mathcal{X}$ subject to the constraints

$$\begin{aligned} u(t) &= x(0) + \int_0^t v(s)\, ds \qquad \forall t \in [0,T] && (5.40) \\ \gamma &= x(0) + \int_0^T v(t)\, dt && (5.41) \\ w(t) &= \xi(t)\,[c(t-\Delta(t)) - x_0] + x_0 + \int_0^{\rho(t)} v(s)[1 - \dot{\Delta}(t)]\, ds \quad \forall t \in [0,T] && (5.42) \end{aligned}$$

Here $\xi(\cdot)$ is defined as in equation (5.38), and $\rho(t) = \max\{0, t - \Delta(t)\}$. Indeed, (5.40) says that $x(t) := u(t)$ is absolutely continuous, (5.41) implies that γ is the endpoint $x(T)$, and (5.42) implies that $w(t) = x(t - \Delta(t))$ for $t \in [0,T]$. The following function $\mathcal{D} : \mathcal{X} \to I\!R$ is used to monitor how far an element $\big(\gamma, u(\cdot), w(\cdot), v(\cdot)\big) \in \mathcal{X}$ is from satisfying (5.40)-(5.42): $\mathcal{D}\big(\gamma, u(\cdot), w(\cdot), v(\cdot)\big)$

$$\begin{aligned} &= \left|\gamma - c(0) - \int_0^T v(t)\, dt\right| + \int_0^T \left|u(t) - c(0) - \int_0^t v(s)\, ds\right|^2 dt \\ &\qquad + \int_0^T \left|w(t) - \xi(t)\,[c(t - \Delta(t) - c(0)] - c(0) - \int_0^{\rho(t)} v(s)\, ds\right|^2 dt. \end{aligned}$$

It is clear that (5.40)-(5.42) hold if and only if $\mathcal{D}\big(\gamma, u(\cdot), w(\cdot), v(\cdot)\big) = 0$. The decoupling principle is contained in the following theorem.

Theorem 5.3.1. *Suppose (H1)-(H3) hold and $\epsilon > 0$. Then there exist a constant $\sigma > 0$, an element $\big(\bar{\gamma}, \bar{u}(\cdot), \bar{w}(\cdot), \bar{v}(\cdot)\big) \in \mathcal{X}$, and absolutely continuous arcs $p(\cdot)$ and $q(\cdot)$ defined on $[0, T]$ so that*

(a) $\mathcal{D}\big(\bar{\gamma}, \bar{u}(\cdot), \bar{w}(\cdot), \bar{v}(\cdot)\big) < \epsilon$;

(b) $\Gamma\big(\bar{\gamma}, \bar{u}(\cdot), \bar{w}(\cdot), \bar{v}(\cdot)\big)$ *is within ϵ of the minimum value in (4.4);*

(c) for almost all $t \in [0, T]$, the map

$$(u, w, v) \mapsto L(t, u, w, v) - \langle \dot{p}(t), u \rangle - \langle \dot{q}(t), w \rangle - \langle p(t) + q(t + \Delta(t)[1 - \dot{\Delta}(t)]), v \rangle + \sigma \Big\{ |u - \bar{u}(t)|^2 + |w - \bar{w}(t)|^2 + |v - \bar{v}(t)|^2 \Big\}$$

is minimized at $(u, w, v) = \big(\bar{u}(t), \bar{w}(t), \bar{v}(t)\big)$; and

(d) the map

$$\gamma \mapsto \ell(\gamma) + \langle \gamma, p(T) \rangle + \sigma |\gamma - \bar{\gamma}|^2$$

is minimized at $\gamma = \bar{\gamma}$.

Proof. Denote by $(P_{\eta,\alpha,\beta})$ the problem of minimizing over all arcs $x(\cdot)$ the functional $\Lambda_{\eta,\alpha(\cdot),\beta(\cdot)}\big(x(\cdot)\big) :=$

$$\ell\big(x(T) + \eta\big) \; + \; \int_0^T \; L\big(t, x(t) + \alpha(t), x(t - \Delta(t)) + \beta(t), \dot{x}(t)\big)\, dt, \tag{5.43}$$

where $x(t)$ is set equal to $c(t)$ for $t \in [-\Delta_0, 0], \big(\eta, \alpha(\cdot), \beta(\cdot)\big) \in I\!R^n \times L^2[0, T] \times L^2[0, T]$, and where $\beta(\cdot) \in L^2[0, T]$ is set to 0 when $t - \Delta(t) < 0$ to ensure that we do not perturb the tail. Note that the integral in (5.43) is well-defined since it is a normal integrand ([25]).

Define a value function $V : I\!R^n \times L^2[0, T] \times L^2[0, T] \to (-\infty, \infty]$ by setting $V\big(\eta, \alpha(\cdot), \beta(\cdot)\big)$ as the optimal value in (5.43). Here $V\big(\eta, \alpha(\cdot), \beta(\cdot)\big) = \infty$ if there are no feasible arcs for $P_{\eta,\alpha,\beta}$. We begin with a lemma.

Lemma 5.3.2. *V is lower semicontinuous. If $V\big(\eta, \alpha(\cdot), \beta(\cdot)\big) < \infty$ then a solution to $P_{\eta,\alpha,\beta}$ exists.*

Proof. This follows from Lemma 5.1.2 in the previous section. $\square$

Let $\epsilon > 0$. The Density Theorem ([3], Theorem 3.1) of proximal analysis guarantees the existence of $\big(\bar{\eta}, \bar{\alpha}(\cdot), \bar{\beta}(\cdot)\big)$ and $\big(\zeta, \phi(\cdot), \psi(\cdot)\big)$ in $\in I\!R^n \times L^2[0,T] \times L^2[0,T]$ satisfying

- $|\bar{\eta}| + \|\bar{\alpha}\|_2 + \|\bar{\beta}\|_2 < \epsilon$
- $|V\big(\bar{\eta}, \bar{\alpha}(\cdot), \bar{\beta}(\cdot)\big) - V(0,0,0)| < \epsilon,$
- $\big(\zeta, \phi(\cdot), \psi(\cdot)\big) \in \partial_P V\big(\bar{\eta}, \bar{\alpha}(\cdot), \bar{\beta}(\cdot)\big)$, where $\partial_P V$ denotes the proximal subgradient of V.

Note that $\psi(t) = 0$ for all $t \in [0, \Delta_0]$ satisfying $t - \Delta(t) < 0$. The proximal subgradient inequality is

$$V(\eta,\alpha,\beta) - V(\bar{\eta},\bar{\alpha},\bar{\beta}) + \sigma'\{|\eta-\bar{\eta}|^2 + \|\alpha-\bar{\alpha}\|_2^2 + \|\beta-\bar{\beta}\|_2^2\}$$
$$\geq \langle \zeta, \eta-\bar{\eta}\rangle + \langle \phi, \alpha-\bar{\alpha}\rangle + \langle \psi, \beta-\bar{\beta}\rangle, \quad (5.44)$$

where $\delta' > 0$ and $\sigma' > 0$ and (5.44) holds for all $(\eta,\alpha,\beta) \in B((\bar{\eta},\bar{\alpha},\bar{\beta}),\delta')$. Let $\bar{x}(\cdot)$ be an optimal solution to $(P_{\bar{\eta},\bar{\alpha},\bar{\beta}})$, which means that $V(\bar{\eta},\bar{\alpha},\bar{\beta}) = \Lambda_{\bar{\eta},\bar{\alpha},\bar{\beta}}(\bar{x})$. It follows from the definition of V, that for any arc $x(\cdot)$ one has $V(\eta,\alpha,\beta) \leq \Lambda_{\eta,\alpha,\beta}(x)$. We may substitute this in (5.44) to obtain:

$$\Lambda_{\eta,\alpha,\beta}(x) - \langle \zeta,\eta\rangle - \langle \phi,\alpha\rangle - \langle \psi,\beta\rangle$$
$$+\,\sigma\{|\bar{\eta}-\eta|^2 + \|\bar{\alpha}-\alpha\|_2^2 + \|\bar{\beta}-\beta\|_2^2\} \geq \text{ibid}\ (\bar{\eta},\bar{\alpha},\bar{\beta},\bar{x}) \quad (5.45)$$

Our notation follows [2], where ibid $(\bar{\eta},\bar{\alpha},\bar{\beta},\bar{x})$ represents the left hand side of the inequality but with the variables $(\bar{\eta},\bar{\alpha},\bar{\beta},\bar{x})$ substituted in place of (η,α,β,x).

The following change of notation converts the parameter variables to perturbations of arcs:

$$
\begin{aligned}
x(\cdot) + \alpha(\cdot) &:= u(\cdot) & \dot{x}(\cdot) &:= v(\cdot) \\
x(T) + \eta &:= \gamma & & \\
\bar{x}(\cdot) + \bar{\alpha}(\cdot) &:= \bar{u}(\cdot) & \dot{\bar{x}}(\cdot) &:= \bar{v}(\cdot) \\
\bar{x}(T) + \bar{\eta} &:= \bar{\gamma} & & \\
x(\cdot - \Delta(\cdot)) + \beta(\cdot) &:= w(\cdot) & & \\
\bar{x}(\cdot - \Delta(\cdot)) + \bar{\beta}(\cdot) &:= \bar{w}(\cdot) & &
\end{aligned}
$$

where $v(t-\Delta(t))$ is set equal to 0 when $t-\Delta(t) \in [-\Delta_0, 0]$. The inequality (5.45) in the new variables becomes

$$
\begin{aligned}
\ell\big(\gamma\big) \;+\; &\int_0^T L\big(t, u(t), w(t), v(t)\big)\, dt - \langle \zeta, \gamma - x(T) \rangle \\
&- \int_0^T \big\{ \langle \phi(t), u(t) - x(t) \rangle + \langle \psi(t), w(t) - x(t-\Delta(t)) \rangle \big\}\, dt \\
&+ \sigma' \Big\{ \|u - x - \bar{u} - \bar{x}\|_2^2 + |\gamma - x(T) - \bar{\gamma} + \bar{x}(T)|^2 + \|w - x - \bar{w} + \bar{x}\|_2^2 \Big\} \\
&\geq \text{ibid } (\bar{\gamma}, \bar{u}, \bar{x}, \bar{w}, \bar{v}).
\end{aligned} \tag{5.46}
$$

Since $2a^2 + 2b^2 \geq (a+b)^2$ for any a, b, the square terms in (5.46) can be divided, and the result is

$$
\begin{aligned}
\ell\big(\gamma\big) \;+\; &\int_0^T L\big(t, u(t), w(t), v(t)\big)\, dt - \langle \zeta, \gamma - x(T) \rangle \\
&- \int_0^T \big\{ \langle \phi(t), u(t) - x(t) \rangle + \langle \psi(t), w(t) - x(t-\Delta(t)) \rangle \big\}\, dt \\
&+ 2\sigma' \Big\{ \|u - \bar{u}\|_2^2 + \|x - \bar{x}\|_2^2 + \|x_\Delta - \bar{x}_\Delta\|_2^2 \\
&\qquad + |\gamma - \bar{\gamma}|^2 + |x(T) - \bar{x}(T)|^2 + \|w - \bar{w}\|_2^2 \Big\} \\
&\geq \text{ibid } (\bar{\gamma}, \bar{u}, \bar{x}, \bar{w}, \bar{v}).
\end{aligned} \tag{5.47}
$$

Here $\|x_\Delta - \bar{x}_\Delta\|_2^2$ is used to denote $\|x(t-\Delta(t)) - \bar{x}(t-\Delta(t))\|_2^2$. The following estimates are used to eliminate the dependence on $x(\cdot)$ terms.

$$\begin{aligned}
\|x-\bar{x}\|_2 &= \left\|(x(0)-\bar{x}(0)) + \int_0^t \{v(s)-\bar{v}(s)\}\,ds\right\|_2 \le \sqrt{T}\,\|v-\bar{v}\|_1 \\
|x(T)-\bar{x}(T)| &= \left|\int_0^T \{v(s)-\bar{v}(s)\}ds\right| \le \|v-\bar{v}\|_1 \\
\|x_\Delta - \bar{x}_\Delta\|_2 &= \left\{\int_0^T \left|\int_0^{\rho(t)} \{v(s)-\bar{v}(s)\}\,ds\right|^2 dt\right\}^{\frac{1}{2}} \le \sqrt{T}(1+K_\Delta)\,\|v-\bar{v}\|_1\,.
\end{aligned}$$

In particular, then

$$(T+1)(1+K_\Delta)\|v-\bar{v}\|_1^2 \ge \max\big\{\|x-\bar{x}\|_2^2,\, |x(T)-\bar{x}(T)|^2,\, \|x_\Delta - \bar{x}_\Delta\|_2^2\big\}, \quad (5.48)$$

and substituting (5.48) into (5.47) yields

$$\begin{aligned}
\ell\big(\gamma\big) - \langle\zeta, \gamma - x(T)\rangle \;+ &\int_0^T \;L\big(t, u(t), w(t), v(t)\big)\,dt \\
&- \int_0^T \{\langle\phi(t), u(t)-x(t)\rangle + \langle\psi(t), w(t) - x(t-\Delta(t))\rangle\}\,dt \\
&+ \sigma\Big\{\|u-\bar{u}\|_2^2 + \|w-\bar{w}\|_2^2 + \|v-\bar{v}\|_1^2 + |\gamma-\bar{\gamma}|^2\Big\} \\
&\ge \text{ibid }(\bar{\gamma}, \bar{u}, \bar{w}, \bar{v}), \qquad (5.49)
\end{aligned}$$

where $\sigma := 6\sigma'(T+1)(1+K_\Delta)$.

Now define the arcs $p(\cdot)$ and $q(\cdot)$ by

$$\begin{aligned}
p(t) &= -\zeta - \int_t^T \phi(s)ds \\
q(t) &= -\int_t^T \psi(s)ds\,.
\end{aligned}$$

which belong to $L^2[0,T]$. The arc $q(\cdot)$ is extended to belong to $L^2[0, T+\Delta_0]$ by defining $q(s) = 0$ for all $s \in [T, T+\Delta_0]$. The next steps are integration by parts:

$$\begin{aligned}
\int_0^T \langle\phi(t), x(t)\rangle dt &= \int_0^T \langle\dot{p}(t), x(t)\rangle\,dt = \langle p(t), x(t)\rangle\Big|_0^T - \int_0^T \langle p(t), v(t)\rangle\,dt \\
&= \langle -\zeta, x(T)\rangle - \langle p(0), x(0)\rangle - \int_0^T \langle p(t), v(t)\rangle dt,
\end{aligned}$$

and

$\int_0^T \langle \psi(t), x(t-\Delta(t))\rangle\, dt$

$$\begin{aligned}
&= \langle q(t), x(t-\Delta(t))\rangle\Big|_0^T - \textstyle\int_0^T \langle q(t), v(t-\Delta(t)[1-\dot{\Delta}(t)])\rangle\, dt \\
&= -\langle q(0), x(-\Delta(0))\rangle - \textstyle\int_0^T \langle q(t)[1-\dot{\Delta}(t)], v(t-\Delta(t))\rangle\, dt \\
&= -\langle q(0), c(-\Delta(0))\rangle - \textstyle\int_0^T \langle q(t+\Delta(t))[1-\dot{\Delta}(t)], v(t)\rangle\, dt,
\end{aligned}$$

since $x(-\Delta(0)) = c(-\Delta(0))$ and $q(s) = 0$ for $s \in [T, T+\Delta_0]$. We now substitute into (5.49) to obtain

$$\begin{aligned}
\ell(\gamma) \; &+ \langle p(T), \gamma\rangle + \sigma|\gamma-\bar{\gamma}|^2 \; + \int_0^T \Big\{ L(t, u(t), w(t), v(t)) \\
&- \langle \phi(t), u(t)\rangle - \langle \psi(t), w(t)\rangle - \langle p(t) + q(t+\Delta(t))[1-\dot{\Delta}(t)], v(t)\rangle \Big\}\, dt \\
&+ \sigma\{\|u-\bar{u}\|_2^2 + \|w-\bar{w}\|_2^2 + \|v-\bar{v}\|_1^2\} \\
&\geq \text{ibid}\,(\bar{\gamma}, \bar{u}, \bar{w}, \bar{v})\,. \qquad (5.50)
\end{aligned}$$

The terms $\langle q(0), c(-\Delta(0))\rangle$, and $\langle p(0), x(0)\rangle$ do not depend on γ, u, w or v, thus, they are subtracted out and need not appear in the inequality (5.50).

Recall that (5.50) holds whenever $\|\alpha-\bar{\alpha}\|_2$, $|\eta-\bar{\eta}|$ and $\|\beta-\bar{\beta}\|_2$ are each less than δ', and by definition

$$\begin{aligned}
\|\alpha-\bar{\alpha}\|_2 &= \|(u-\bar{u}) - (x-\bar{x})\|_2, \\
|\eta-\bar{\eta}| &= |(\gamma-\bar{\gamma}) - (x(T)-\bar{x}(T))|, \qquad \text{and} \\
\|\beta-\bar{\beta}\|_2 &= \|(w-\bar{w}) - (x(t-\Delta(t)) - \bar{x}(t-\Delta(t))\|_2.
\end{aligned}$$

Now (5.48) says that each of $\|x-\bar{x}\|_2$, $\|x_\Delta - \bar{x}_\Delta\|_2$, and $|x(T)-\bar{x}(T)|$ are small when $\|v-\bar{v}\|_1$ is small, and it is also clear that $\|w-\bar{w}\|_2$ is small if $\|u-\bar{u}\|_2$ is small. Therefore (5.50) holds whenever $\|u-\bar{u}\|_2, \|v-\bar{v}\|_1$, and $|\gamma-\bar{\gamma}|$ are small.

We now verify that statements (a) through (d) of the theorem hold. (a) We must show $\mathcal{D}\big(\bar{\gamma}, \bar{u}(\cdot), \bar{w}(\cdot), \bar{v}(\cdot)\big) < \epsilon$. Notice that

$$\int_0^T \left| \bar{u}(t) - x(0) - \int_0^t \bar{v}(s)\, ds \right|^2 dt = \|\bar{u} - \bar{x}\|_2 = \|\bar{\alpha}\|_2^2,$$

$$\left| \bar{\gamma} - x(0) - \int_0^T \bar{v}(t)\, dt \right| = |\bar{\gamma} - \bar{x}(T)| = |\bar{\eta}|,$$

which provide bounds for the first two terms in the definition of $\mathcal{D}$. For the third term, observe that

$$\begin{aligned}
&\left| \bar{w}(t) - \xi(t)\left[c(t - \Delta(t)) - x_0\right] - c(0) - \int_0^{\rho(t)} \bar{v}(s)\, ds \right|^2 dt \\
&\qquad = \left| \bar{w}(t) - c(0) - \int_0^{t-\Delta(t)} \bar{v}(s)\, ds \right|^2 \\
&\qquad = \left|\bar{\beta}(t)\right|^2,
\end{aligned}$$

which implies the third term is bounded by $\|\bar{\beta}\|_2^2$. But, $\|\bar{\alpha}\|_2^2 + |\bar{\eta}| + \|\bar{\beta}\|_2^2 < \|\bar{\alpha}\|_2 + |\bar{\eta}| + \|\bar{\beta}\|_2 < \epsilon$ by our original choice of $\bar{\alpha}$, $\bar{\eta}$, and $\bar{\beta}$, so that statement (a) is satisfied.

(b) This is clear because $\Gamma\big(\bar{\gamma}, \bar{u}(\cdot), \bar{w}(\cdot), \bar{v}(\cdot)\big) = V(\bar{\eta}, \bar{\alpha}, \bar{\beta})$ and $V(0,0,0)$ is the minimum for equation (4.4). Therefore, since $|V(\bar{\eta}, \bar{\alpha}, \bar{\beta}) - V(0,0,0)| < \epsilon$, it follows that $\Gamma\big(\bar{\gamma}, \bar{u}(\cdot), \bar{w}(\cdot), \bar{v}(\cdot)\big)$ is within ϵ of the minimum value in (4.4) and (b) holds.

(c) The goal is to convert the integral inequality in (5.50) into the pointwise statement

$$\begin{aligned}
(u, w, v) \mapsto L(t, u, w, v) - \langle \dot{p}(t), u\rangle - \langle \dot{q}(t), w\rangle - \langle p(t) + q(t + \Delta(t))[1 - \dot{\Delta}(t)], v\rangle \\
+ \sigma\left\{ |u - \bar{u}(t)|^2 + |w - \bar{w}(t)|^2 + |v - \bar{v}(t)|^2 \right\}
\end{aligned}$$

is minimized at $(u, w, v) = \big(\bar{u}(t), \bar{w}(t), \bar{v}(t)\big)$ for almost all $t \in [0, T]$.

Let f_t denote the function

$$f_t(u,w,v) := L(t,u,w,v) - \langle \dot{p}(t), u\rangle - \langle \dot{q}(t), w\rangle - \langle p(t) + q(t+\Delta(t))[1-\dot{\Delta}(t)], v\rangle$$
$$+ \sigma\{|u-\bar{u}(t)|^2 + |w-\bar{w}(t)|^2 + |v-\bar{v}(t)|^2\}. \quad (5.51)$$

Hypotheses (H2) and (H3) imply that f_t attains a minimum for almost every t, and hence it is sufficient to establish that

$$A(r) := \left\{ t \in [0,T] : f_t(\bar{u}(t), \bar{w}(t), \bar{v}(t)) \geq \min_{(u,w,v)\in R^{3n}} f_t(u,w,v) + r \right\}$$

has measure zero for every $r > 0$.

Suppose not, and so there exists an $r > 0$ with $m(A(r)) > 0$. Let B_i be a decreasing sequence of subsets of $A(r)$ such that for each i one has

$$\frac{1}{i} m(A(r)) < m(B_i) < \frac{3}{i} m(A(r)).$$

Let $(u'(\cdot), w'(\cdot), v'(\cdot))$ be measurable such that $(u'(t), w'(t), v'(t))$ minimizes f_t for almost each t, whose existence is guaranteed by standard measurable selection results [25].

We next show $u'(\cdot)$,$w'(\cdot)$, and $v'(\cdot) - \bar{v}(\cdot)$ belong to $L^2[0,T]$ (notice that $\bar{v}(\cdot)$ is only in $L^1[0,T]$). Consider the inequality

$$f_t(u'(t), w'(t), v'(t)) \leq f_t(\bar{u}(t), \bar{w}(t), \bar{v}(t)),$$

which, after simplifying and inserting the definition (5.51), becomes

$$L(t,u'(t),w'(t),v'(t)) - \langle \dot{p}(t), u'(t)\rangle - \langle \dot{q}(t), w'(t)\rangle + \langle p(t) + q(t+\Delta(t))[1-\dot{\Delta}(t)], v'(t)\rangle$$
$$+ \sigma\{|u'(t)-\bar{u}(t)|^2 + |w'(t)-\bar{w}(t)|^2 + |v'(t)-\bar{v}(t)|^2\}$$
$$\geq \text{ibid}\left(\bar{u}(t), \bar{w}(t), \bar{v}(t)\right).$$

Denoting the lower bound of L by b, and rearranging terms, we have

$$\sigma\{|u'(t)-\bar{u}(t)|^2+|w'(t)-\bar{w}(t)|^2+|v'(t)-\bar{v}(t)|^2\}$$
$$\leq \langle \dot{p}(t), u'(t)-\bar{u}(t)\rangle+\langle \dot{q}(t), w'(t)-\bar{w}(t)\rangle+\langle p(t)+q(t+\Delta(t))[1-\dot{\Delta}(t)], v'(t)-\bar{v}(t)\rangle$$
$$+L(t,\bar{u}(t),\bar{w}(t),\bar{v}(t))-b\,. \quad (5.52)$$

Set

$$\begin{aligned} W(t) &= (u'(t)-\bar{u}(t), w'(t)-\bar{w}(t), v'(t)-\bar{v}(t))\ , \\ g(t) &= \Big(\dot{p}(t),\dot{q}(t),p(t)+q(t+\Delta(t))[1-\dot{\Delta}(t)]\Big)\ , \\ k(t) &= L(t,\bar{u}(t),\bar{w}(t),\bar{v}(t))-b\,. \end{aligned}$$

Then inequality (5.52) becomes $\sigma\,|W|^2-\langle g,W\rangle-k\leq 0$ and by the quadratic formula

$$|W|^2 \leq \frac{|g|+\sqrt{|g|^2+4k\sigma}}{2\sigma}.$$

Thus, $(2\sigma\,|W|^2-|g|)^2\leq |g|^2+4k\sigma$. Since $g\in L^2[0,T]$ and $k\in L^1[0,T]$ it follows that the integral of $(2\sigma\,|W|^2-|g|)^2$ is bounded and thus, $W\in L^2[0,T]$, proving that u', w',and $v'-\bar{v}$ are in $L^2[0,T]$. Now, since $v'-\bar{v}\in L^2[0,T]$ and $\bar{v}\in L^1[0,T]$, it follows that v' is also in $L^1[0,T]$.

We continue now with the proof of (c). Define a family of functions $(u_i,w_i,v_i)\in L^2[0,T]\times L^2[0,T]\times L^1[0,T]$ as:

$$(u_i(t),w_i(t),v_i(t))=\begin{cases}(u'(t),w'(t),v'(t)), & \text{if } t\in B_i;\\ (\bar{u}(t),\bar{w}(t),\bar{v}(t)), & \text{otherwise}\,.\end{cases} \quad (5.53)$$

Then,

$$\lim_{i\to\infty}\|(u_i,w_i,v_i)-(\bar{u},\bar{w},\bar{v})\|_2=0.$$

From the definition of (u',w',v') and since $m(A(r))>0$, it follows that

$$\int_0^T f_t(u_i,w_i,v_i)\,dt<\int_0^T f_t(\bar{u},\bar{w},\bar{v})\,dt.$$

But setting $\gamma = \bar{\gamma}, u = u_i, w = w_i$, and $v = v_i$ for i sufficiently large, this contradicts inequality (5.50). Hence, it must be that $m(A(r)) = 0$, which proves (c) and completes our proof.

(d). This is verified by setting $u = \bar{u}, v = \bar{v}, w = \bar{w}$ in (5.50). $\square$

As an immediate consequence of statement Theorem 5.3.1(c) is that the inclusion

$$\big(\dot{p}(t), \dot{q}(t), p(t) + q(t + \Delta(t))[1 - \dot{\Delta}(t)]\big) \in \partial_P L\big(t, \cdot, \cdot, \cdot)(\bar{u}(t), \bar{w}(t), \bar{v}(t)\big) \tag{5.54}$$

holds for almost all $t \in [0, T]$, which is a rudimentary form of the Euler-Lagrange equation. Here $\partial_P L(t, \cdot, \cdot, \cdot)(\bar{u}(t), \bar{w}(t), \bar{v}(t)\big)$ refers to the proximal subgradient of L with respect to the (x, y, v) variables evaluated at $(\bar{u}(t), \bar{w}(t), \bar{v}(t)\big)$ (see [3]). An immediate consequence of (d) is that

$$-p(T) \in \partial_P \ell(\bar{\gamma}), \tag{5.55}$$

which is a forerunner of the transversality condition.

The inclusion (5.54) will next be dualized and caste in Hamiltonian form. Define $H_\sigma : [0, T] \times I\!R^n \times I\!R^n \times I\!R^n \to I\!R$ by

$$H_\sigma(t, x, y, p) := \sup_{v \in R^n} \Big\{ \langle p, v \rangle - L(t, x, y, v) - \sigma |v - \bar{v}(t)|^2 \Big\}.$$

The case $\sigma = 0$ is the usual Hamiltonian and is denoted simply as H.

Proposition 5.3.3. *With the assumptions and notation of the previous theorem, the following hold. Throughout we define* $\bar{q}(t - \Delta(t)) = q(t) - \Delta(t)[1 - \dot{\Delta}(t)]$

(i) $\dot{p}(t) \in \partial_P(-H_\sigma)\big(t, \cdot, \bar{w}(t), p(t) + \bar{q}(t + \Delta(t))\big)\big(\bar{u}(t)\big)$

(ii) $\dot{q}(t) \in \partial_P(-H_\sigma)\big(t, \bar{u}(t), \cdot, p(t) + \bar{q}(t + \Delta(t))\big)\big(\bar{w}(t)\big)$

(iii) $\bar{v}(t) \in \partial_P H\big(t, \bar{u}(t), \bar{w}(t), \cdot\big)\big(p(t) + \bar{q}(t + \Delta(t))\big)$

If H is in addition locally Lipschitz in (x, y, p) then

(iv) $(-\dot{p}(t), -\dot{q}(t), \bar{v}(t)) \in \partial_C H\big(t, \cdot, \cdot, \cdot\big)(\bar{u}(t), \bar{w}(t), p(t) + \bar{q}(t + \Delta(t))),$

also holds, where ∂_C denotes the Clarke generalized gradient. In the above inclusions, the subgradients are taken with respect to the $(\cdot)$ variable.

Proof. Assertions (i)-(iii) will follow directly from Theorem 5.3.1(c), where we let $t \in [0, T]$ be so that

$$\begin{aligned} L(t, u, w, v) - \langle \dot{p}(t), u\rangle - \langle \dot{q}(t), w\rangle - \langle p(t) + \bar{q}(t + \Delta(t)), v\rangle& \\ + \sigma\Big\{ |u - \bar{u}(t)|^2 + |w - \bar{w}(t)|^2 + |v - \bar{v}(t)|^2 \Big\}& \\ \geq \text{ibid } (t, \bar{u}(t), \bar{w}(t), \bar{v}(t))& \end{aligned} \quad (5.56)$$

holds for all (u, w, v) near $(\bar{u}(t), \bar{w}(t), \bar{v}(t))$. In fact (5.56) holds whenever (u, w) is near $(\bar{u}(t), \bar{w}(t))$ and for any $v \in I\!\!R^n$, since the left hand side is convex in v. A rearrangement of terms gives

$$\begin{aligned} \sup_{v\in R^n} &\left\{\langle p(t) + \bar{q}(t + \Delta(t)), v\rangle - L(t, u, w, v) - \sigma\, |v - \bar{v}(t)|^2\right\} \\ &\leq \langle \dot{p}(t), \bar{u}(t) - u\rangle + \langle \dot{q}(t), \bar{w}(t) - w\rangle + \sigma \left\{|u - \bar{u}(t)|^2 + |w - \bar{w}(t)|^2\right\} \\ &\qquad + \langle p(t) + \bar{q}(t + \Delta(t)), \bar{v}(t)\rangle - L(t, \bar{u}(t), \bar{w}(t), \bar{v}(t)), \end{aligned}$$

which is equivalent to

$$\begin{aligned} (-H_\sigma)(t, u, w, p(t) + \bar{q}(t + \Delta(t))) &\geq (-H_\sigma)(t, \bar{u}(t), \bar{w}(t), p(t) + \bar{q}(t + \Delta(t))) \\ &\quad + \langle \dot{p}(t), u - \bar{u}(t)\rangle + \langle \dot{q}(t), w - \bar{w}(t)\rangle \\ &\quad - \sigma \left\{|u - \bar{u}(t)|^2 + |w - \bar{w}(t)|^2\right\}. \end{aligned} \quad (5.57)$$

Both (i) and (ii) are implied by (5.57). The inequality (5.56) with $u = \bar{u}(t)$ and $w = \bar{w}(t)$ is

$$L(t, \bar{u}(t), \bar{w}(t), v) - \langle p(t) + \bar{q}(t + \Delta(t)), v\rangle + \sigma\,|v - \bar{v}(t)|^2$$
$$\geq L(t, \bar{u}(t), \bar{w}(t), \bar{v}(t)) - \langle p(t) + \bar{q}(t + \Delta(t)), \bar{v}(t)\rangle,$$

and thus

$$p(t) + \bar{q}(t + \Delta(t)) \in \partial_P L(t, \bar{u}(t), \bar{w}(t), \cdot)(\bar{v}(t)). \tag{5.58}$$

But $L(t, \bar{u}(t), \bar{w}(t), \cdot)$ is convex, and so the convex and proximal subgradients coincide. Hence (iii) holds from (5.58) and the duality relationship between subgradients of convex conjugates.

To prove (iv), suppose that H is locally Lipschitz in (x, y, p), from which it is clear that H_σ is also locally Lipschitz. Thus, for any $(\mu_1, \mu_2, \mu_3) \in I\!R^{3n}$, $H_\sigma^\circ\big(t, \bar{u}(t), \bar{w}(t), p(t) + \bar{q}(t + \Delta(t)); (\mu_1, \mu_2, \mu_3)\big)$ majorizes

$$\limsup_{\lambda\to 0} \frac{1}{\lambda} \cdot \{\mathrm{H}_\sigma(t, \bar{u}(t), \bar{w}(t), p(t) + \bar{q}(t + \Delta(t)) + \lambda\mu_3)$$
$$- H_\sigma(t, \bar{u}(t) - \lambda\mu_1, \bar{w}(t) - \lambda\mu_2, p(t) + \bar{q}(t + \Delta(t)))\}$$
$$\geq \limsup_{\lambda\to 0} \frac{1}{\lambda} \cdot \{\mathrm{H}_\sigma(t, \bar{u}(t), \bar{w}(t), p(t) + \bar{q}(t + \Delta(t))) + \lambda\langle\bar{v}(t), \mu_3\rangle$$
$$- H_\sigma(t, \bar{u}(t) - \lambda\mu_1, \bar{w}(t) - \lambda\mu_2, p(t) + \bar{q}(t + \Delta(t)))\}$$
$$\geq \limsup_{\lambda\to 0} \frac{1}{\lambda} \cdot \left\{\langle\dot{p}(t), -\lambda\mu_1\rangle + \langle\dot{q}(t), -\lambda\mu_2\rangle - \sigma\left\{|\lambda\mu_1|^2 + |\lambda\mu_2|^2\right\} + \lambda\langle\bar{v}(t), \mu_3\rangle\right\}$$
$$= -\langle\dot{p}(t), \mu_1\rangle - \langle\dot{q}(t), \mu_2\rangle + \langle\bar{v}(t), \mu_3\rangle.$$

The last inequality is justified via (5.57). Thus the inclusion in (iv) is proved in case where H is replaced by H_σ. However, by a key technical result of Loewen and Rockafellar [13], one has $\big(-\dot{p}(t), -\dot{q}(t), \bar{v}(t)\big) \in \partial_C H$ as well, which finishes the proof of (iv). $\square$

5.4 Necessary Conditions for the Delayed Problem of Bolza (DPB)

There is a large literature of necessary conditions for nonsmooth problems without time delay, e.g. [5, 13, 14, 24, 31] and their cited references. The recent monograph [4] further extends and unifies this subject. References with time delay include [7, 6, 16]. Here we shall only show the result in [2] can be extended to varying time delay under the following additional hypotheses:

(H1') $\ell(\cdot)$ is locally Lipschitz, and

(H4) $L(t, \cdot, \cdot, \cdot)$ is locally Lipschitz and satisfies the growth conditions

$$|\zeta + \xi| \leq k_1(t)|\theta| + k_2(t) \qquad\qquad |\xi| \leq k_3(t)$$

for almost all $t \in [0, T]$ and all $(\zeta, \xi, \theta) \in \partial_P L(t, x, y, v)$, where $k_1(\cdot)$ belongs to $L^2[0, T]$, and $k_2(\cdot)$ and $k_3(\cdot)$ belong to $L^1[0, T]$.

Theorem 5.4.1. *Assume (H1'), (H2)-(H4), and suppose $x(\cdot)$ solves the Bolza problem (4.4). Then there exist arcs $p(\cdot)$ and $q(\cdot)$ with $\bar{q}(\cdot - \Delta(\cdot)) = (\cdot - \Delta(\cdot))[1 - \dot{\Delta}(t)]$ satisfying the following:*

(a) For almost every $t \in [0, T]$,

$$\big(\dot{p}(t), \dot{q}(t), p(t) + \bar{q}(t + \Delta(t))\big) \in \partial_C L\big(t, x(t), x(t - \Delta(t)), \dot{x}(t)\big),$$

(b) for almost every $t \in [0, T]$,

$$\big(-\dot{p}(t), -\dot{q}(t), \dot{x}(t)\big) \in \partial_C H\big(t, x(t), x(t - \Delta(t)), p(t) + \bar{q}(t + \Delta(t))\big),$$

and

(c) $-p(T) \in \partial_L \ell(x(T))$ (where ∂_L is the limiting subgradient).

Proof. It will later be desirable for $x(\cdot)$ to be the *unique* solution to (4.4), and this simplification is effectively achieved by modifying L as

$$\widetilde{L}(t,x,y,v) := L(t,x,y,v) + |v-\dot{x}|^2.$$

Theorem 5.3.1 and Proposition 5.1.3 is applied to this data for a sequence $\epsilon_i \searrow 0$ as $i \to \infty$. There exist $(\gamma_i, \bar{u}_i(\cdot, w_i(\cdot), v_i(\cdot) \in \mathcal{X}$ and arcs $p_i(\cdot)$ and $q_i(\cdot)$ satisfying

(i) For almost every $t \in [0,T]$,

$$(\dot{p}_i(t), \dot{q}_i(t), p_i(t) + \bar{q}_i(t+\Delta(t))) \in \partial_P \widetilde{L}(t, \bar{u}_i(t), \bar{w}_i(t), \bar{v}_i(t)),$$

(ii) for almost every $t \in [0,T]$,

$$(-\dot{p}_i(t), -\dot{q}_i(t), \bar{v}_i(t)) \in \partial_C \widetilde{H}(t, \bar{u}_i(t), \bar{w}_i(t), p_i(t) + \bar{q}_i(t+\Delta(t))),$$

and

(iii) $-p_i(T) \in \partial_P \ell(\bar{\gamma}_i)$

Indeed, (i) follows from Theorem 5.3.1(c), (ii) from Proposition 5.1.3 (since the modified Hamiltonian $\widetilde{H}_\sigma$ associated to $\widetilde{L}$ is locally Lipschitz in (x,y,p) by (H4)), and (iii) from Theorem 5.3.1(d). The rest of the proof is a somewhat standard limiting argument.

Since $x(\cdot)$ solves (4.4) and L is locally Lipschitz, Theorem 5.3.1(b) implies that $\|\bar{v}_i(\cdot) - \dot{x}(\cdot)\|_2 \to 0$ (even though $\dot{x}(\cdot)$ itself need not belong to $L^2[0,T]$). This in conjunction with Theorem 5.3.1(a) implies $\gamma_i \to x(T)$ and both $\bar{u}_i(\cdot) \to x(\cdot)$, $\bar{w}_i(\cdot) \to x(\cdot - \Delta(\cdot))$ in $L^2[0,T]$ as $i \to \infty$. It then follows that (an unrelabeled subsequence) satisfies

$$L(t, \bar{u}_i(t), \bar{w}_i(t), \bar{v}_i(t)) \to L(t, x(t), x(t-\Delta(t)), \dot{x}(t)) \tag{5.59}$$

for almost all $t \in [0,T]$ as $i \to \infty$. The local Lipschitz condition (H1') on $\ell(\cdot)$, that $\{\gamma\}_i$ is bounded, and (iii) imply $\{p_i(T)\}_{i=1}^{\infty}$ is also bounded, and therefore has a convergent subsequence (which again is not relabeled). The limit $p(T)$ will be the endpoint of and arc $p(\cdot)$, and belongs to $\partial_L \ell(x(T))$, proving (c). We have to show taking limits lead to (a) and (b) as well.

Using the growth hypotheses in (H4) and (i), we have that

$$\begin{aligned}
\big|\dot{p}_i(t) + \dot{q}_i(t+\Delta(t))\big| &\leq |p_i(T)| + \int_{t+\Delta(t)}^{t} |\dot{q}_i(s)|\,ds + \int_t^T |\dot{p}_i(s) + \dot{q}_i(s)|\,ds \\
&\leq |p_i(T)| + \|k_3\|_1 + \int_t^T \Big\{ k_1(s)\big|p_i(s) + q_i(s)\big| + 2|\bar{v}_i(t) - \dot{x}(t)| + k_2(s) \Big\}\,ds \\
&\leq \bar{\alpha} + \int_t^T \Big\{ k_1(s)\big|p_i(s) + q_i(s+\Delta(s))\big| \Big\}\,ds,
\end{aligned}$$

where $\bar{\alpha} \geq \big|p_i(T)\big| + (1+K_\Delta)\|k3\|_1 + \sqrt{T}\|\bar{v}_i - \dot{x}\|_2 \Big\} + \|k_2\|_1$ is a constant independent of i. Grownwall's inequality implies that $\{p_i(\cdot) + q_i(\cdot)\}_i$ is a sequence uniformly bounded on $[0,T]$, and therefore (H4) implies $\dot{p}_i(t) + \dot{q}_i(t)$ is bounded almost everywhere by an integrable function. Thus by familiar compactness arguments, there exists a subsequence of $\{\dot{p}_i(\cdot) + \dot{q}_i(\cdot)\}$ that converges weakly. By the second growth assumption in (H4), a subsequence of $\{\dot{q}(\cdot)\}$ also converges weakly, which subsequently implies a subsequence of $\{\dot{p}_i(\cdot)\}$ converges weakly. We do not relabel all these subsequences. Furthermore, the arcs $\{p_i(\cdot)\}$, $\{q_i(\cdot)\}$ converge uniformly to the limiting arcs $p(\cdot)$, $q(\cdot)$, and as known (see [3]), the limiting arcs satisfy the limiting version of (i) and (ii). These equations have the form

($\tilde{a}$) $(\dot{p}(t), \dot{q}(t), p(t) + \bar{q}(t+\Delta(t))) \in \partial_C \widetilde{L}(t, \bar{x}(t), \bar{x}(t-\Delta(t)), \dot{x}(t))$ a.e.

($\tilde{b}$) $(-\dot{p}(t), -\dot{q}(t), \dot{x}(t)) \in \partial_C \widetilde{H}(t, \bar{x}(t), \bar{x}(t-\Delta(t)), p(t))$ a.e.

Finally, the results quoted earlier by Loewen and Rockafellar [13] permit the removal of the tilde from ($\tilde{a}$) and ($\tilde{b}$), and thus completes the proof of the theorem. $\square$

5.5 Necessary Conditions for the Neutral Problem of Bolza vs. Necessary Conditions for the Delayed Problem of Bolza

It is worth noting the relationship between the results obtained in the neutral case and the results for the case with delay occurring only in the state variable [20]. The main difference occurs in statement (c) of the main decoupling Theorem. In the state delay case one had the inner product $\langle p(t) + q(t + \Delta(t))[1 - \dot{\Delta}(t)], v(t)\rangle$, whereas in the neutral case, one has two inner products $\langle p(t), v_1(t)\rangle$, and $\langle q(t)[1 - \dot{\Delta}(t)], v_2(t)\rangle$. Given that $v_2(t)$ is defined as $v_1(t - \Delta(t))$, we can see that

$$\langle q(t)\rho(t), v_2(t)\rangle = \langle q(t)\rho(t), v_1(t - \Delta(t))\rangle = \langle q(t + \Delta(t))\rho(t), v_1(t)\rangle,$$

where $\rho(t) = [1 - \dot{\Delta}(t)]$ and the last equality follows from a change of variable. Therefore, our two inner products can be combined to give $\langle p(t) + q(t + \Delta(t))\rho(t), v_1(t)\rangle$, which is then equivalent to the inner product for the state delay case. Part (c) of the decoupling theorem is used to derive the Euler-Lagrange inclusion. Thus, the incorporation of the variable v_2 gives rise to the separation in the relationship between the arcs p and q to the subdifferential of L with respect to the $x(\cdot)$ and $\dot{x}(\cdot)$ variables respectively. In the case of delay in the state variable only, our results resemble those obtained by Mordukhovich [16]. However, in the neutral case, the author is not aware of any results for which the arcs p and q were included in the Euler-Lagrange inclusion separately as in Theorem 5.2.1, part (1). Even in the case of the necessary conditions established in the calculus of variations, [12], p and q appear as terms related by addition.

Chapter 6
Weak Invariance and Time Delay

In this chapter we derive the existence of Euler solutions for time delayed differential equations and characterize weak invariance properties for differential inclusions with time delay in terms of the lower hamiltonian. A characterization of weak invariance properties for these delayed differential inclusions in terms of the Bouligand tangent cone was obtained previously by Haddad [11].

Let $F(x, y)$ be an autonomous multifunction satisfying $F : I\!\!R^n \times I\!\!R^n \rightrightarrows I\!\!R^n$, where the double arrow denotes a map into subsets of $I\!\!R^n$, with the following properties:

(a) For every $(x, y) \in I\!\!R^n \times I\!\!R^n$, $F(x, y)$ is a non empty, compact, convex set.

(b) F is upper semicontinuous

(c) For some γ, β, c greater than zero and for all $(x, y) \in I\!\!R^n \times I\!\!R^n$

$$v \in F(x, y) \Rightarrow \|v\| \leq \gamma\|x\| + \beta\|y\| + c. \tag{6.1}$$

We will consider the differential inclusion: $\dot{x}(t) \in F(x(t), y(t))$ a.e. $t \in [0, T]$ with $y(t) = x(t - \Delta)$, where $x(s) = \phi(s)$ a.e. for $s < 0$ and $\phi : [-\Delta, 0] \rightarrow I\!\!R^n$ is absolutely continuous. First, we will show that an Euler solution can be found for a given selection of the inclusion. Second, we characterize weak invariance for the system (S, F) in terms of the lower hamiltonian $h(x, y, p)$. The lower hamiltonian $h : I\!\!R^n \times I\!\!R^n \times I\!\!R^n \rightarrow I\!\!R$ is defined as

$$h(x, y, p) = \min_{v \in F(x,y)} \langle p, v \rangle.$$

We now provide two results that are of importance in what follows. The proofs may be found in *Nonsmooth Analysis and Control Theory*, [3]. Lemma 6.0.1 may be found in Chapter 4, p.179 and Lemma 6.0.1 may be found in Chapter 3, p. 164.

Lemma 6.0.1. *Let x be an arc on $[a,b]$ which satisfies*

$$\|\dot{x}(t)\| \leq \gamma\|x(t)\| + c(t) \;\; a.e., t \in [a,b]$$

where $\gamma \geq 0$ and $c(\cdot) \in L^1[a,b]$. Then, for all $t \in [a,b]$, we have

$$\|x(t) - x(a)\| \leq (e^{\gamma(t-a)} - 1)\|x(a)\| + \int_a^t e^{\gamma(t-s)}c(s)\,ds.$$

Proposition 6.0.2. *Let $\{v_i\}$ be a sequence in $L_n^2[a,b]$ such that*

$$v_i(t) \in F(\tau_i(t), u_i(t)) + r_i(t)\bar{B} \;\; a.e.,\; t \in [a,b],$$

where F is a multifunction with the above growth hypotheses and where the sequence of measurable functions $\{\tau_i(\cdot), u_i(\cdot)\}$ converges a.e. to $(t, u_0(t))$, and the nonnegative measurable functions $\{r_i\}$ converge to 0 in $L^2[a,b]$. Then there exists a subsequence (we do not relabel), $\{v_i\}$ which converges weakly in $L_n^2[a,b]$ to a limit $v_0(\cdot)$ which satisfies

$$v_0(t) \in F(t, u_0(t)) \;\; a.e.,\; t \in [a,b].$$

We refer to Lemma 6.0.1 as Grownwall's Lemma and to Proposition 6.0.2 as the Weak Sequential Compactness Theorem.

6.1 Euler Solutions for Differential Equations with Delay

Euler solutions for delay differential equations can be shown to exist by applying minor modifications to the method found in *Nonsmooth Analysis and Control Theory* ([3], pp. 180-185) for differential equations without time delay. We proceed to extend this method ([3], p. 181) for equations with constant delay $\Delta > 0$.

Consider the initial value problem:

$$\dot{x}(t) = f(t, x(t), x(t - \Delta)) \ , \ x(s) = \phi(s) \ , \ s \in [-\Delta, 0], \tag{6.2}$$

where ϕ is a uniformly continuous function, $\phi : [-\Delta, 0] \to I\!R^n$, and define $x(0) = \phi(0) = x_0$. Here $f : [0, T] \times I\!R^n \times I\!R^n \to I\!R^n$ and satisfies the growth condition $\|f(t, x, y)\| \leq \gamma\|x\| + \beta\|y\| + c$. We begin by discretizing the time. Let $\pi = \{t_0, t_1, \ldots, t_N\}$ be a, not necessarily uniform, partition of $[0, T]$, where $t_0 = 0 \leq t_1 \leq \ldots \leq t_N = T$. We proceed with the Euler method in the usual way. Consider the differential equation on the interval $[t_0, t_1]$ with constant right-hand side

$$\dot{x}(t) = f(t_0, x_0, y_0), \ x(t_0) = x_0$$

where $y_0 = x(t_0 - \Delta) = x(-\Delta) = \phi(-\Delta)$. This equation has a unique solution $x(t)$ on $[t_0, t_1]$. Using this solution to define $x_1 := x(t_1)$, we may now consider on $[t_1, t_2]$ the initial value problem

$$\dot{x}(t) = f(t_1, x_1, y_1), \ x(t_1) = x_1$$

where $y_1 = x(t_1 - \Delta)$. If $t_1 - \Delta \leq t_0$ we use the function ϕ to find this value. Otherwise, it must be that $t_0 < t_1 - \Delta \leq t_1$ and from the previous step we may use the solution found for the interval $[t_0, t_1]$ to find this value. Again a unique solution exits. The only possible problem that we need to consider is if the value of x at $t_1 - \Delta$ is somehow undefined then we should just replace y_1 with x_0 and in general $y_n = x_{n-1}$.

We can proceed similarly until we have an arc x_π, which is piecewise affine and defined on $[0, T]$. Notice that this arc depends on the partition π defined. Let μ_π be the diameter of the partition π, i.e. $\mu_\pi := \max\{t_i - t_{i-1}, 1 \leq i \leq N\}$.

Define an *Euler solution* for the differential equation with delay to be any arc x which is the uniform limit of Euler polygonal arcs x_{π_j}, corresponding to some sequence π_j converging to zero, which denotes $\mu_{\pi_j} \to 0$. As in the book [3], we may call this solution an Euler arc for f.

In addition, this Euler solution is Lipschitz, as in the case without time delay. We state the results from our discussion above more formally in the following theorem.

Theorem 6.1.1. *Suppose that the function f satisfies the linear growth condition*

$$\|f(t,x,y)\| \leq \gamma\|x\| + \beta\|y\| + c$$

for all $(t,x,y) \in [a,b] \times I\!R^n \times I\!R^n$, and where γ, β, and c are positive. Then at least one Euler solution x to the initial-value problem (6.2) exists on $[a,b]$, and any Euler solution is Lipschitz.

6.2 Weak Invariance for Delay Inclusions

A system (S,F) consisting of a closed set S and a multifunction F mapping $I\!R^n \times I\!R^n$ to nonempty, compact, convex subsets of $I\!R^n$, is said to be *weakly invariant* provided that for all $x_0 \in S$, there is a trajectory x satisfying the differential inclusion $\dot{x}(t) \in F(x(t), x(t-\Delta))$ for $t \in [0,\infty)$ so that

$$x(0) = x_0, \qquad x(t) \in S \quad \forall t \geq 0.$$

In order to characterize weak invariance for delayed differential inclusions we provide some technical results that allow us to pass from Euler solutions to trajectories of our multifunctions and prove the invariance property for a given function that satisfies the growth condition (6.1), which is satisfied by each element in $F(x,y)$. These results are an extension of those obtained in the case without time delay in ([3],pp. 188-192). These technical results allow us to prove our main result, the characterization of weak invariance for the delayed differential inclusion in terms of the lower hamiltonian.

Lemma 6.2.1. *Let $\{x_i\}$ be a sequence of arcs on $[-\Delta, b]$ such that $x_i(s) = \phi(s)$ a.e. for each $s < 0$, where $\phi(\cdot)$ is an absolutely continuous function defined on $[-\Delta, 0]$. Also, $\{x_i(a)\}$ is bounded and for each i, x_i satisfies*

$$\dot{x}_i(t) \in F(\tau_i(t), x_i(t) + y_i(t), x_i(t-\Delta) + y_i(t-\Delta)) + r_i(t)B \text{ a.e.} \tag{6.3}$$

where $y_i(s) = 0$ when $s < 0$ and $\{y_i\}, \{r_i\}, \{\tau_i\}$ are sequences of measurable functions on $[a, b]$ such that y_i converges to 0 in L^2, $r_i \geq 0$ converges to 0 in L^2, and τ_i converges a.e. to t. Then, there is a subsequence of $\{x_i\}$ that converges uniformly to an arc x, which is a trajectory of F and whose derivative converges weakly to $\dot{x}$.

Proof. From the differential inclusion and the linear growth condition we have

$$\begin{aligned} \|\dot{x}_i(t)\| &\leq \gamma\|x_i(t) + y_i(t)\| + \beta\|x_i(t-\Delta) + y_i(t-\Delta)\| + |r_i(t)| \qquad (6.4)\\ &\leq 2K\|x_i(t) + y_i(t)\| + |r_i(t)| \qquad (6.5)\end{aligned}$$

where $K = \max\{\gamma, \beta\}$. From Gronwall's Lemma

$$\|x_i(t) - x_i(a)\| \leq (\mathrm{e}^{2K(t-a)} - 1)\|x_i(a)\| + \int_a^t \mathrm{e}^{2K(t-s)}|r_i(s)|\, ds$$

so that

$$\begin{aligned} \|x_i(t)\| &\leq \|x_i(a)\| + (\mathrm{e}^{2K(t-a)} - 1)\|x_i(a)\| + \int_a^t \mathrm{e}^{2K(t-s)}|r_i(s)|\, ds \\ &\leq \mathrm{e}^{2K(t-a)}\|x_i(a)\| + \mathrm{e}^{2K(b-a)}\|r_i(s)\|_1. \qquad (6.6)\end{aligned}$$

Thus, x_i is uniformly bounded for each i. Given this bound for x_i and (6.4), its derivative $\dot{x}_i$ is bounded in L^2. Hence, there exists a subsequence $\{\dot{x}_i\}$, and we do not relabel, that converges weakly to $v_0 \in L^2[a, b]$ with $\{x_i\}$ converging uniformly to x. Passing to the limit in

$$x_i(t) = x_i(a) + \int_a^t \dot{x}_i(s)\, ds$$

we have

$$x(t) = x(a) + \int_a^t v_0(s)\, ds$$

so that x is an arc and $\dot{x} = v_0$. Applying the Weak Sequential Compactness Theorem and applying the limit to (6.3) we obtain

$$\dot{x}(t) \in F(t, x(t), x(t - \Delta)),$$

which completes our proof. □

Corollary 6.2.2. *Let f be any selection of F and let x be an Euler solution on $[a, b]$ of $\dot{x}(t) = f(t, x(t), x(t-\Delta))$, $x(a) = x_0$, and $x(s) = \phi(s)$ for each $s \in [-\Delta, 0]$. Then, x is a trajectory of F on $[a, b]$*

Proof. Let x_{π_j} be the polygonal arcs whose uniform limit is x. Let $t \in (a, b)$ be a non-partition point and let $\tau_j(t)$ be the partition point t_i immediately before t. Then,

$$\begin{aligned} \dot{x}_{\pi_j}(t) = f(t_i, x_i(t), x_i(t - \Delta)) &\in F(t_i, x_i(t), x_i(t - \Delta)) \\ &= F(\tau_j(t), x_{\pi_j}(t) + y_i(t), x_{\pi_j}(t - \Delta) + y_i(t - \Delta)) \end{aligned} \quad (6.7)$$

where $y_i(t) = x_i(t) - x_{\pi_j}(t) = x_{\pi_j}(\tau_j(t)) - x_{\pi_j}(t)$. Since the functions x_{π_j} admit a common Lipschitz constant K, we have that

$$\|y_j(t)\|_\infty \le \sup_{t \in [a,b]} |\tau_j(t) - t| \le K\mu_{\pi_j}$$

and in particular,

$$\|y_j(t - \Delta)\|_\infty \le \sup_{t \in [a,b]} |\tau_j(t) - \Delta - (t - \Delta)| \le K\mu_{\pi_j}.$$

Therefore, y_j is uniformly bounded in $[-\Delta, b]$. It follows that τ_j and y_j are measurable and converge uniformly to t and 0 respectively. Taking the limit in (6.7), using the above Lemma, shows that the uniform limit x of x_{π_j} is a trajectory of F. □

Finally, we proceed to prove the invariance property for a given function that satisfies the growth condition (6.1).

Lemma 6.2.3. *Let S be a closed set. Let f satisfy $\|f(t,x,y)\| \leq \gamma\|x\|+\beta\|y\|+c$ for all $(t,x,y) \in [0,T]\times I\!R^n \times I\!R^n$. Let $x(\cdot)$ be an Euler arc for f on $[0,T]$. Let Ω be an open set containing $x(t)$ for all $t \in [-\Delta, T]$ and suppose that every $(z,w) \in \Omega\times\Omega$ with $w(t) \in S$ for $t \in [-\Delta, 0]$ satisfies the following: There exists $s \in proj_S(z)$ such that $\langle f(t,z,w), z-s\rangle \leq 0$. Then we have $d_S(x(t)) \leq d_S(x(0))$, for all $t \in [0,T]$.*

Proof. Let x_π be a polygonal arc in the sequence converging uniformly to x. Let $\pi = \{t_0,\ldots,t_N\}$ be a partition of $[0,T]$ and denote the node at t_i by x_i for $(i = 0,1,\ldots N)$ so that $x_0 = x(0)$ and y_0 is the value of $x(t-\Delta)$ on this interval. Suppose $x_\pi \in \Omega$ for $t \in [0,T]$. Choose $s_i \in proj_S(x_i)$ such that $\langle f(t,x_i,w_i), x_i - s_i\rangle \leq 0$. Choose k such that $\|\dot{x}_\pi\|_\infty < k$ then,

$$\begin{aligned}
{d_S}^2(x_1) &\leq \|x_1 - s_0\|^2 \\
&= \|x_i - x_0\|^2 + \|x_0 - s_0\|^2 + 2\langle x_1 - x_0, x_0 - s_0\rangle \\
&\leq k^2(t_1-t_0)^2 + d_S(x_0)^2 + 2\int_{t_0}^{t_1} \langle \dot{x}_\pi(t), x_0 - s_0\rangle dt \\
&\leq k^2(t_1-t_0)^2 + d_S(x_0)^2 + 2\int_{t_0}^{t_1} \langle f(t_0,x_0,y_0), x_0 - s_0\rangle dt \\
&\leq k^2(t_1-t_0)^2 + d_S(x_0)^2.
\end{aligned}$$

Notice that the last inequality follows since $\langle f(t_0,x_0,y_0), x_0 - s_0\rangle$ is non-negative. The same estimate can be applied to any node x_i to obtain,

$$\begin{aligned}
{d_S}^2(x_i) &\leq {d_S}^2(x_{i-1}) + k^2(t_i - t_{i-1})^2 \\
&\leq {d_S}^2(x_0) + k^2\sum_{\alpha=1}^{i}(t_\alpha - t_{\alpha-1})^2 \\
&\leq {d_S}^2(x_0) + k^2\mu_\pi\sum_{\alpha=1}^{i}(t_\alpha - t_{\alpha-1}) \\
&\leq {d_S}^2(x_0) + k^2\mu_\pi(b-a).
\end{aligned}$$

We may apply this estimate to the family of Euler arcs x_{π_j} by replacing x_i with x_{π_j} and μ_π with μ_{π_j} in the above estimate. Since $\mu_{\pi_j} \to 0$, passing to the limit we obtain

$$d_S(x(t)) \leq d_S(x(a))$$

for each $t \in [a, b]$. $\square$

Given the above results, we can now prove that weak invariance for delay inclusions can be characterized in terms of the *lower Hamiltonian* h corresponding to $F(x, y)$.

Theorem 6.2.4. *Let S be a closed set. If for all $x \in S$, $h(x, y, N_S^P(x)) \leq 0$, then there exists a solution to $\dot{x}(t) \in F(x(t), x(t - \Delta))$ with $x(t) \in S$ for all $t \geq 0$. Here $x(s) = \phi(s)$ for all $s \in [-\Delta, 0]$, ϕ is an absolutely continuous function on $[-\Delta, 0]$, and $\phi(s) \in S$ for all $s \in [-\Delta, 0]$.*

Proof. Let $x \in \mathbf{R}_n$. Chose $s \in proj_S(x)$, $y \in I\!R^n$, and $s_2 \in proj_S(y)$. Choose $v \in F(s, s_2)$ which minimizes $v \to \langle v, x - s\rangle$. Let $f_p(x, y) = v$. Since $x - s \in N_S^P(s)$, $\langle f_p(x, y), x - s\rangle \leq 0$. Also,

$$\begin{aligned}
\|f_p(x, y)\| &= \|v\| \leq \gamma\|s\| + \beta\|s_2\| + c \\
&\leq \gamma\|s - x\| + \gamma\|x\| + \beta\|s_2 - y\| + \beta\|y\| + c \\
&\leq \gamma d_S(x) + \gamma\|x\| + \beta d_S(y) + \beta\|y\| + c \\
&\leq \gamma\|x - s_0\| + \|\gamma\|x\| + \beta\|y - s_0\| + \beta\|y\| + c \\
&\leq 2\gamma\|x\| + 2\beta\|y\| + (\gamma + \beta)\|s_0\| + c.
\end{aligned}$$

Therefore, f_p satisfies the growth condition (6.1). By the above Lemma 6.2.3, for any $x_0 \in S$, the Euler solutions to $\dot{x} = f_p(x, y)$, with $x(0) = x_0$ and $y(s) \in S$ for $s \in [-\Delta, 0]$, lie in S. Now, we must show that x is a trajectory of F. Define

$F_s(x,y) := \text{co}\,\{F(s,s_2) : s \in proj_S(x), s_2 \in proj_S(y)\}$. It can be shown that F_s satisfies the standing hypotheses and that $F_s(x,y) = F(x,y)$ if $x,y \in S$. By definition, $f_p \in F_s(x,y)$. So, the Euler solution to $\dot{x} = f_p(x,y)$, say x, is a trajectory for F_s a.e. on $[0,T]$. And, since $x,y \in S$, we have that $F_s(x,y) = F(x,y)$ and thus, x is a trajectory of F. $\square$

In the nondelay case, there are further characterizations of weak invariance in terms of tangential cones. In particular, weak invariance of a nondelay system (S,F) has been shown to be equivalent to our characterization in terms of the lower hamiltonian and also to Haddad's characterization in terms of the Bouligand tangent cone ([3],p. 193). We expect that the same equivalence follows in the case of time delay.

References

[1] F. H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley & Sons, New York, 1983.

[2] F. H. Clarke, A decoupling principle in the calculus of variations, *J. of Mathematical Analysis and Applications*, **172** (1993), 92-105.

[3] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, P. R. Wolenski, *Nonsmooth Analysis and Control Theory*, Springer, New York, 1998.

[4] F. H. Clarke, Necessary conditions in dynamic optimization, (2002), preprint.

[5] F.H. Clarke, *Methods of Dynamic and Nonsmooth Optimization*, CBMS-NSF Regional Conf. Series 57, SIAM, Philadelphia, 1989.

[6] F. H. Clarke, P. R. Wolenski, The sensitivity of optimal control problems to time delay, *SIAM J. Control Optim.*, **29** (1991), 1176-1215.

[7] F. H. Clarke, G. G. Watkins, Necessary conditions, controllability and the value function for differential-difference inclusions, *J. of Nonlinear Analysys, Theory, Methods and Applications*, **10** (1986), 1155-1179.

[8] R.E. Edwards, *Functional Analysis*, Dover, New York, 1965.

[9] L.È. Èl'sgol'c, *Qualitative Methods in Mathematical Analysis*, AMS, Translations of Mathematical Monographs, **12**, Rhode Island, 1968.

[10] H. Gorceki, S. Fuksa, P. Grabowski, A. Korytowski, *Analysis and Synthesis of Time Delay Systems*, John Wiley & Sons, 1989.

[11] G. Haddad, Monotone trajectories of differential inclusions and functional differential inclusions with memory, *Israel J. of Mathematics*, **39** (1981), 83-100.

[12] D.K. Hughes, Variational and optimal control problems with delayed argument, *J. of Optim. Theory and Applications*, **2** (1968), 1-14.

[13] P. D. Loewen and R. T. Rockafellar, The adjoint arc in nonsmooth optimization, *Trans. Amer. Math. Soc.*, **325** (1991), 39-72.

[14] P. D. Loewen and R. T. Rockafellar, New necessary conditions for the generalized problem of Bolza, *SIAM J. Control Optim.*, **34** (1996), 1496-1511.

[15] A. S. Matveev, The instability of optimal control problems to time delay, accepted (2004).

[16] B. S. Mordukhovich and R. Trubnik, Stability of discrete approximations and necessary optimality conditions for delay-differential inclusions. Optimization with data perturbations, II, *Ann. Oper. Res.*, **101** (2001), 149-170.

[17] B.S. Mordukhovich and L. Wang, Optimal control of hereditary differential inclusions, *Proceedings of the 41st joint IEEE-CDC Conference*, Las Vegas (2002).

[18] B.S. Mordukhovich and L. Wang, Optimal control of neutral functional-differential inclusions, *SIAM J. Control Optim.*, **43** (2004), 111-136.

[19] N. Ortiz, Necessary conditions for the neutral problem of Bolza with continuously varying time delay, *J. of Mathematical Analysis and Applications*, accepted (2004).

[20] N. Ortiz and P. R. Wolenski, Decoupling Time-Delays, *Proceedings of the 41st joint IEEE-CDC Conference*, Las Vegas (2002).

[21] N.L. Ortiz, P.R. Wolenski, The decoupling technique for continuously varying time delay systems, *Set-Valued Analysis*, **00** (2004), 1-15.

[22] N.L. Ortiz, P.R. Wolenski, An existence theorem for neutral variational problems of Bolza, *J. of Mathematical Analysis and Applications*, **289** (2004), 260-265.

[23] L. Rifford, Existence of Lipschitz and semiconcave control-Lyapunov functions, *SIAM J. Control Optim.*, **39** (2000), 1043-1064.

[24] R. T. Rockafellar, Equivalent subgradient versions of Hamiltonian and Euler-Lagrange equations in variational analysis, *SIAM J. Control Optim.*, **34** (1996), 1300-1314.

[25] R. T. Rockafellar, *Nonlinear Operators and the Calculus of Variations*, Springer Verlag Lecture Notes in Mathematics, **543** (1976), 157-207.

[26] R. T. Rockafellar, Integral functionals, normal integrands, and measurable selections, In:L. Waelbroeck (ed.), *Nonlinear Operators and the Calculus of Variations*, Lecture Notes in Math. 543, Springer, Berlin, 1976, 157-207.

[27] R.T. Rockafellar, Integrals which are convex functionals II, *Pacific J. Math.*, **39** (1971), 439-469.

[28] R.T. Rockafellar, Existence theorems for general control problems of Bolza and Lagrange, *Adv. in Math.*, **15** (1975), 312-333.

[29] R.T. Rockafellar, Optimal arcs and the minimal time function in problems Lagrange, *Trans. Amer. Math. Soc.*, **180** (1973), 53-84.

[30] L. D. Sabbagh, Variational problems with lags, *J. of Optim. Theory and Applications*, **3** (1969), 34-51.

[31] R. Vinter, *Optimal Control*, Birkhâuser, Boston, 2000.

Vita

Norma Leticia Ortiz was born on February 18 1972, in San Juan, Puerto Rico. She finished her undergraduate studies at the University of Louisana, Lafayette, in May 1996. She earned a master of science degree in mathematics from Louisiana State University in May 2001. In January 2000, she came to Louisiana State University to pursue graduate studies in mathematics. She is currently a candidate for the degree of Doctor of Philosophy in mathematics, which will be awarded in May 2005.